12-2017

Modeling effective work groups and teams : an assessment of the inclusion of social value orientation.

Gary Daniel Jones
University of Louisville

MODELING EFFECTIVE WORK GROUPS AND TEAMS: AN ASSESSMENT OF
THE INCLUSION OF SOCIAL VALUE ORIENTATION

By

Gary Daniel Jones
B.S., United States Military Academy at West Point, 2004
M.S., University of Delaware, 2013

A Dissertation
Submitted to the Faculty of the
College of Education & Human Development at the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy
in Educational Leadership and Organizational Development

Department of Educational Leadership, Evaluation and Organizational Development
University of Louisville
Louisville, Kentucky

December 2017

MODELING EFFECTIVE WORK GROUPS AND TEAMS: AN ASSESSMENT OF
THE INCLUSION OF SOCIAL VALUE ORIENTATION

By

Gary Daniel Jones
B.S., United States Military Academy at West Point, 2004
M.S., University of Delaware, 2013

A Dissertation Approved on

October 31, 2017

by the following Dissertation Committee:

_____
Dr. Denise Cumberland

_____
Dr. Meera Alagaraja
Co-Chair

_____
Dr. Jeffrey Valentine

_____
Dr. Richard Balkin

DEDICATION

Throughout my professional career, I have been blessed to find inspiration, challenge, and achievement. Meaningful work and committed coworkers have helped me find fulfillment and happiness through vocation. This dissertation is dedicated to the hope that it may, in some small part, help others to be so lucky.

ACKNOWLEDGEMENTS

Dr. Mike Benver – thank you for being an excellent partner in the workplace and for setting me on this path towards a doctoral degree.

The officers, civilian professionals, and recruiters of the Third Recruiting Brigade – the time I spent on your team were some of my favorite years in the Army. I wish you all the best in the future. Colonels McTigue and Hertel – thank you both for believing in me and supporting me, allowing me to do my best to improve the unit.

My dissertation committee – I have been both challenged and educated by all of you, each in your own way. The four of you are consummate professionals and I could not have had a better group to work with and learn from.

Binah – you have been so gracious and patient. Thank you for understanding.

ABSTRACT

MODELING EFFECTIVE WORK GROUPS AND TEAMS: AN ASSESSMENT OF
THE INCLUSION OF SOCIAL VALUE ORIENTATION

Gary Daniel Jones

October 31, 2017

The purpose of this study is to examine research on effective work groups in terms of measurable input characteristics and relevant collective outputs. Antecedents and consequences of group processes are explored; models of work groups and teams are examined and assessed in light of subsequent research. Additionally, social value orientation – a behavioral trait known to predictably influence interpersonal outcomes – is introduced as a concept that is also relevant to group work. A revised model that includes social value orientation as a predictive factor for group productivity is presented and tested. This empirical study is a correlative, quantitative investigation of extant work groups and teams within the United States Army Recruiting Command. Primary statistical tools are multilevel modeling and Markov Chain Monte Carlo simulations. Key results include identifying group potency as the most predictive variable of performance and providing evidence that social value orientation is significantly related to group productivity over and above other included variables. Implications and discussion of relevance to human resource development (HRD) are included.

Keywords: Work Groups and Teams, Workplace Productivity, Social Value Orientation, HRD

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Work groups and teams are an important part of human resource development (Raes, Kyndt, Decuyper, Van Den Bossche, & Dochy, 2015) and organizations today fully recognize that effective teams are integral to achieving strategic business goals and objectives (Gilley & Kerno, 2010; Venneberg, 2010). Even though HRD scholars study work groups, examining various dimensions of diversity within teams (Jehn, Chadwick, & Thatcher, 1997; Pelled, Eisenhardt, & Xin, 1999), intra-group conflict in the workplace (O'Connor, Gruenfeld, & McGrath, 1993), and the effect of managerial support on employee satisfaction (Miles & Mangold, 2002); existing research and theory on this topic is somewhat limited. Given the importance of work groups in this highly global economy, human resource development (HRD) scholars need to ensure that workplace-relevant theories are routinely updated, refined, and applied. Concerns over the relevance of theory and practical utility are of primary importance to the HRD community and one of four current unresolved HRD issues (Russ-Eft, 2016). Do we have the best model to assess antecedents to work group productivity? How can we explain differences between work groups than an organization assesses as meeting or failing to meet their expected achievements? This study seeks to answer this question by examining the pioneering work of Campion, Medsker, and Higgs (1993) and the revised work of Campion, Papper,

1

and Medsker (1996) that has been used to conceptualize and measure work group productivity, as well as subsequent relevant research.

Campion et al. (1993; 1996) approached this problem by aggregating smaller models of various collective processes and functions into a cohesive and comprehensive scheme of correlated inputs and outputs. This approach differs distinctly from some other conceptualizations, which focus more on explaining the inter-workings of teams and are less interested in predictive inputs and their relation to desired outputs (such as Hinsz, Tindale, & Vollrath, 1997). The Campion Work Group Effectiveness Model (CWGEM) is explanatory and predictive, connecting inputs and outputs of group processes together in a cohesive framework. While the CWGEM aggregates and consolidates extensive research fields into a consolidated model and is theoretically compelling, it suffers from some limitations. The instruments associated with the model are difficult to use in a real-world setting, limiting utility for HRD practitioners, and Campion et al.'s empirical assessments of the entire model (1993; 1996) are somewhat unconvincing due to methodological concerns.

Additionally, this study introduces Social Value Orientation (SVO) and suggests its potential utility as a predictor of workplace outcomes. Social Value Orientation is a personality trait that categorizes individuals based on preferences in the balance of outcomes between oneself and others in interpersonal situations. This trait (SVO) has been shown to be predictive of outcomes in HRD applications such as organizational climate (Bogaert, Boone, & van Witteloostuijn, 2012) and employee compensation schemes (Upton, 2009). This trait is lifelong and stable (Knight & Dubro, 1984), and supported by a large body of research in experimental psychology (DeDrue, 2010;

2

McClintock, Messick, Kuhlman, & Campos, 1973; Messick & McClintock, 1968; Van Lange & Liebrand, 1991).

Social Value Orientation is an individual's preference of outcome distribution between themself and others in situations of social dilemma. Social dilemmas are any interpersonal situation where outcome distribution is partially or totally the result of one or more participant's decision (often the resulting distribution of outcomes is the interaction between the decisions of the participating parties), and where incentives of individual actors do not align (de Cremer, Snyder, & DeWitte, 2001). Restated, SVO is an individual's preference for outcome distribution in any social conflict in terms of that which benefits oneself in relation to that which benefits others. Critical differences exist between those that are generally oriented to favor others over themselves (*altruists*) or prefer a balance between outcomes (*cooperators*) - these two orientations, together, can be called *pro-social* - compared to those that are either more self-preferring (*individualists*) or view social dilemmas as competitions to be won (*competitors*) - these two orientations together can be called *pro-self*. Pro-self individuals tend to expect and plan for deception or social rule-breaking from others with whom they deal, and primarily concern themselves with the outcome rather than individual's actions that produced the outcome (Stouten, de Cremer, & van Dijk, 2005). Conversely, pro-social individuals view such violations as injustices and are prone to actions that might redress these wrongs, caring less for overall outcome or group productivity and more about adherence to rules and equity in the groups' processes. Stouten et al. argue that pro-self individuals care more about efficiency and the bottom line, while pro-social individuals care more about the process and adherence to norms.

In small groups, including typical work groups, SVO predicts individual behaviors and interpersonal attitudes (Sonnemas, van Dijk, & van Winden, 2001). Scholars argue that to be in a group with others is to be in conflict (see De Drue & Beersma, 2005, for an extensive review of organizational literature on interpersonal conflict). When tangible and intangible rewards, benefits, and detriments are contingent on outcome of a group, then the situation and context qualifies as a social dilemma to which SVO theory speaks. Specific to HRD, significant concern has been given to performance management of individuals within groups at work, because group members may take or receive credit for outcomes (good and bad) they did not directly create (Siders, George, & Dharwadkar, 2001; Williams & Plouffe, 2007; Zampetakis, 2014) – this is the connection between abstract social dilemmas (and SVO) and practitioner concerns about real-world impacts to the bottom line.

These issues are important because incentives to underperform and to limit member contributions can occur in groups. Working in groups provides a choice to each member (Wagner, 1995): i.e., avoid contributing at all (become a free-rider), put forth less than one would have working as an individual (social loafing), or fully participate. As will be seen, the CWGEM partially considers these factors concerning the relative size of work groups and whether the workload is properly shared amongst group member. However, SVO potentially provides a better stand-in for social loafing or free-riding, as pro-socials resent and police these violations of commitment to social norms and group effort (Stouten et al., 2005).

Also included in this study are a revised model, testable propositions, empirical investigation, and implications for HRD theory and practice. Model revisions take into

4

account subsequent attempts at testing portions of Campion's model while the testable propositions are articulations of how one could assess the suitability of the revised model and of SVO in it. Following a review of relevant literature is a methods section detailing the study. Included are the organizational context of the study, discussions of the population and sample, and conceptual and operationalized definitions of variables. Results and discussion of the empirical study conclude this paper.

This study uses a correlational research design comparing work group-level predictors with work group outcomes. This research differs from extant work group studies in that it is comprehensive (includes a wide array of independent variables) and also compares their relationship to outcome variables using multilevel modeling analysis (instead of the bivariate correlational methods used by Campion et al. 1993). Additionally, because the organization used in this study is hierarchical (i.e. work groups are managed by several levels of management and clustered geographically), multilevel modeling – also called hierarchical linear modeling – is used to identify the echelons in the organization where the factors weighing on work group performance are located. No known studies of work group performance or outcomes use this method. This study provides evidence of the most important differentiators between high-performing and low-performing work groups within the same organization, and speaks to the location (in terms of managerial hierarchy) of those differentiators. To summarize the discussion thus far, this investigation can be restated as two overarching and related research questions:

- Research Question One: Is social value orientation predictive of work group outcomes?

- Research Question Two: How can the model used by Campion et al. (1993; 1996) be improved?

Work Group Classification

Within the context of an organization, a work group (or work team) is a unique subgroup of employees. While there are different schemes to categorize and differentiate groups and teams within larger organizations, common threads emerge. According to Sundstrom, De Meuse, and Futrell (1990), non-virtual work teams have important characteristics: shared work space and purpose, but with varying size, duration (either temporary or enduring), and function (either advice and involvement, or production and service, or projects and development, or action and negotiation). This classification accords with the more rigorous scheme of Cohen and Bailey (1997), who classify groups in organizational settings as either work teams (enduring, output-producing), parallel teams (special-purpose projects teams such as quality-improvement or problem-solving teams), project teams (temporary with a one-time specified outcome), and management teams (semi-permanent teams of leaders collectively overseeing a large subordinate hierarchy).

For this study, the terms *work group* and *work team* will be used interchangeably, and inclusively refers to the work teams of Cohen and Bailey (1997), or to enduring (i.e. non-temporary) versions of Sundstrom et al.'s (1990) production and service teams; unless otherwise noted all discussion of work groups or work teams should be assumed to meet these criteria - this is important because it limits the scope of this study to only these kinds of teams, not other teams found in the workplace such as supervisory oversight committees or quality management groups. Cohen and Bailey's parallel teams, project

teams, and management teams are all excluded from consideration. Work groups and teams are capable of exercising subunit power in competition with other parts of their parent organization (Mintzberg, 1983) and collective achievement is important as work groups preoccupy themselves with successful outcomes to ensure their survival (Ancona, 1990). The work units included in the empirical investigation fully meet these definitions as well.

CHAPTER II

LITERATURE REVIEW

The major focus of this review is to present relevant literature on the consequences and outcomes of effective work groups and teams as well as the antecedents correlated with these desired outcomes. The foundation of this investigation is an assessment of the comprehensive work group effectiveness model proposed and tested by Campion, Medsker, and Higgs (1993) and revised by Campion, Papper, and Medsker (1996). This model (Campion's Work Group Effectiveness Model) is explanatory and predictive, connecting inputs and outputs of group processes together in a cohesive framework. While this work is theoretically compelling and aggregates and consolidates extensive research fields into a consolidated model, Campion's Work Group Effectiveness Model suffers some limitations. Campion et al.'s two empirical assessments of the entire model (1993; 1996) together investigate work groups and teams in a total of five real-world organizations, but are unconvincing and raise some methodological concerns; the instruments associated with the model are difficult to use in a real-world setting, limiting their utility. These concerns are discussed in the following sections.

Structure and Organization of Review

This chapter is divided into three sections. The first section of this paper discusses the Campion Work Group Effectiveness Model (CWGEM). This paper

8

explores the CWGEM primarily through a review of subsequent literature on the components of the model. This is to say, while no other studies are known to have applied the whole model to an organization, many studies apply a portion of the model to a problem and report the results, and these studies are introduced and discussed. After the CWGEM is described and considered in light of more recent research, a brief discussion of methodological and theoretical concerns will conclude the first section.

The second section of this chapter introduces a stable personality trait known to have predictable and testable impacts on group settings: Social Value Orientation (SVO). As an effort to extend Campion's work group effectiveness model, SVO – a behavioral trait known to be an effective predictor of interpersonal outcomes – is considered as an alternative conceptualization of certain portions of the Campion work group effectiveness model. Social Value Orientation has been used in research involving work groups and teams and has been found to be a relevant factor to consider, but has not been included (in all reviewed studies) in a comprehensive model of work group effectiveness.

The third section of this chapter concludes with the proposal of a revised model which incorporates a measure of SVO. This section is primarily a discussion of potential changes or modifications to the model that will address the concerns and shortcomings identified in the first section.

Campion's Work Group Effectiveness Model

The seminal work of Campion et al. (1993) aggregated and extended various work by numerous other scholars across multiple disciplines (examples include Gladstein, 1984 and Hackman, 1987). Primary research lines incorporated into the Campion work group effectiveness model emerged from organizational psychology,

social psychology, industrial engineering, and socio-technical theory. Campion et al. explicitly categorized and consolidated varied group process research and organizational behavior research into five thematic inputs (further divided into a total of 19 subthemes) and three distinct outputs. The three outputs are *productivity* (an objective tally of the tangible work produced by a group), *satisfaction* (subjective opinion of a work group's performance by the constituent members of the work group), and *manager judgement* (subjective opinion of a work group's performance by management). The five input themes and 19 input subthemes are *job design* (subdivided into *participation*, *self-management*, *task variety*, *task identity*, and *task significance*), *interdependence* (*task interdependence*, *goal interdependence*, and *interdependent feedback and rewards*), *context* (*training, managerial support, cooperation and coordination between groups*), *composition* (*heterogeneity, flexibility, relative size, preference for group work*), and *process* (*potency, workload sharing, social support*, and *cooperation and coordination within group*). Figure 1 provides an overview of the relationships between the five input themes, 19 subthemes, and three consequences or outputs of work groups.

Figure 1. Campion Work Group Effectiveness Model.

Themes and Subthemes as Inputs to Group Processes   Work Group Outputs

Job Design	Self-Management
	Participation
	Task Variety
	Task Significance
	Task Identity

Interdependence	Task Interdependence
	Goal Interdependence
	Interdependent Feedback & Rewards

Composition	Heterogeneity
	Flexibility
	Relative Size
	Preference for Group Work

Context	Training
	Managerial Support
	Communication / Cooperation Between Groups

Process	Potency
	Social Support
	Workload Sharing
	Communication / Cooperation Within Groups

Productivity

Employee Satisfaction

Manager Judgement

Figure 1. Depiction of theorized relationships between inputs and outputs of work groups. The 19 input subthemes are grouped into five themes on the left side of the figure. The three outputs are listed on the right side. Adapted from "Relations between work group characteristics and effectiveness: implications for designing effective workgroups," by M.A. Campion, G.J. Medsker, and A.C. Higgs, 1993, *Personnel Psychology*, 46, p. 825.

Campion's Work Group Effectiveness Model is useful because it is applicable to a wide variety of organizational contexts in that it does not rely on predictors that are specific to one type of organization. This is to say, it is generalizable to varied organizations in ways that industry-specific models are not (in stark contrast to a work group effectiveness model such as Piercy, Cravens & Morgan, 1999, which only applies

11

to groups of territorially-bounded salespeople) because it is domain-independent and does not rely on work group predictions that are specific to certain tasks or functions. This model has the potential to be applicable to work groups regardless of context because the inputs and outputs are non-domain specific and the five input themes are conceptually distinct from and independent of each other, allowing for research to use restricted models by investigating only some of the factors in the model.

Initial Empirical Study

In the initial publication of the work group effectiveness model, Campion et al. (1993) included an empirical study to provide evidence for construct validity and predictive validity. The sample for this initial study was 391 employees and 70 managers from 80 work groups in a large financial services company (spread across the U.S.). Custom-created measures for the 19 input subthemes of the model explained 73% of the total variance in differences in work group outputs between the 80 sampled workgroups in the study. Interestingly, the Campion et al. study measured the input sub-themes two different ways: from five group members per group and from one member of management with appropriate knowledge and oversight per group. While many of the subthemes had been tested individually prior to 1993, using them together and testing them concurrently provided evidence of composite predictive validity; these 19 subthemes have predictive capability when measured together.

Additionally, this study provided empirical evidence that individual job-level characteristics that Campion et al. aggregated into group-level themes and subthemes could in fact be used at the group level (i.e. this might be called evidence of aggregation validity; evidence that summing and averaging individual-level constructs to the group

level provides effective predictors of outcomes at the group level). Finally, Campion's

work group effectiveness model as a whole is a thoughtful consideration of the varied

ways that a work group might be successful or productive; these outputs are discussed

next.

Work Group Outputs

Campion et al. (1993) subdivided work group outputs into objective production

and subjective satisfaction based on previous works by Gladstein (1984) and Hackman

(1987). *Productivity* is an objective assessment of the work group's achievement of the

task for which they exist, often involving counting items or tracking achievement of

milestones in projects; *satisfaction* measures the opinions of the group members with

their work group, and is intertwined with perceptions of group identity and group

potential. Campion et al. also consider satisfaction from the perspective of the larger

organization, in terms of *manager judgements* of group performance, which speaks to the

degree in which a work group meets the needs of its organization in subjective terms.

Work Group Inputs

Campion et al. (1993) categorize input factors that impact levels of work group

performance *themes*. These five themes (*job design*, *interdependence*, *composition*,

context, and *process*) are themselves each comprised of multiple concepts relating to the

structure or function or context of a work group. In the following section, each input

theme and associated subthemes are defined, and findings from the initial study are

included. Following a complete description and assessment of the initial study from

Campion et al., subsequent research that speaks to the theme or subtheme is included and

discussed.

Job design. This theme explains the relationship between designs of jobs –
specifically how each group member's job interacts with each other – and group
productivity. Jobs are designed properly when group members participate in decision-
making (*participation,* the first job design subtheme), exercise *self-management* in
determining how to accomplish their tasks, allow each group member to perform multiple
different tasks for the group (*task variety*), perform tasks that group members believe are
important (*task significance*), and the group's outcome is clear and distinct, recognizable
as a result of the efforts of group members (*task identity*).

During their initial 1993 study of group productivity in a financial services
company, Campion et al. found significant correlation between job design subthemes and
productivity. As mentioned, all sub-themes were measured from group members and
from management. When group members assessed these five job design sub-themes, two
were significantly correlated with group productivity: *self-management* and *participation*
significantly positively correlated with group productivity ($r = .23$; $p < .05$ and $r = .18$;
$p < .10$, respectively). When group supervisors made assessments of subordinate groups,
three of the five job design subthemes significantly positively correlated with group
performance: *self-management* ($r = .18$; $p < .10$), *participation* ($r = .22$; $p < .10$), and *task
variety* ($r = .20$, $p < .05$). Neither employee nor managerial assessment of *task
significance* nor *task identity* correlated significantly with group productivity. Theme-
wide composite employee and manager ratings of job design were significantly correlated
with group production at the level for both employee-based ($r = .19$; $p < .05$) and
managerial-based ($r = .25$; $p < .05$) assessments. Subsequent group research by Erez and
Somech (1996) and De Drue and West (2001) generally agree with the theory, while

game theory-based laboratory experiments find that autocratic leadership (associated with low self-management and low participation in decision-making) likely leads to a destabilized group and excessive membership turnover (van Vugh, Jepson, Hartman, & de Cremer, 2004).

Erez and Somech (1996) conducted a study to determine why some groups were less productive than others. This research was conducted with two organizational contexts in Israel: managers from Israeli communal society (kibbutz) and managers from urban organizations. While they did not explicitly address *task identity*, the research design of Erez and Somech assumed that tasks were clearly defined and their task outputs were clearly distinct and attributable to the efforts of the group. This presumption of high levels of *task identity* allowed the researchers to manipulate other factors in their research (such as varied subcultures between the two organizational contexts) to identify conditions that might lead to loss of production in groups (specifically due to social loafing). Because of this, one can infer then that adequate *task identity* is a necessary (though not sufficient) precondition to work group productivity.

In the performance of tasks that were highly identifiable, the work groups studied by Erez and Somech (1996) had varying levels of *participation* and *task importance*. Work groups from the kibbutz had high *participation* (decisions were made collectively) and high *task importance* (apparent because the members of the work groups elected to take on the tasks to accomplish; presumably they would not bother to collectively and publicly commit to things they did not feel was important). No evidence of high levels of *participation* or *task importance* was presented in the urban work groups. In comparison to the groups led by the urban managers, the kibbutz groups had higher group production

(attributed by the authors to the group *participation* and implied *task importance* norms of their kibbutz subculture). Although normed and studied in the United States, the CWGEM is not culturally anchored, i.e. the relationships between inputs and outputs are nominally independent of culture, an assumption that might be untenable upon further inspection. The Erez and Somech research emphasizes the impact that cultures and subcultures might have on the predispositions to certain levels of certain inputs. However, the relationships between inputs and outputs (specifically that under conditions of high task identity, groups with high participation and high task importance are more productive than groups of lower participation and task importance) proposed by Campion et al. (1993), are supported by this study of managers in different Israeli subcultures.

Participation in the decision making process was found to be significantly predictive of team innovation (De Drue and West, 2001), especially when accompanied by high levels of voiced dissent by minority opinions. Two experiments by De Drue and West found small, positively-correlated effects ($\Delta R^2 = .16$ and $.11$, respectively, after controlling for known covariates; $p < .05$ for both) between participation levels and a desired work group outcome (team innovation) - De Drue and West used the three-item measure for participation developed by Campion et al. (1993) and further tested by Campion et al. (1996), providing evidence for internal validity (reporting Cronbach's α of .83 and .85 for two samples). While this study does not speak directly to the predictive power of participation in terms of production, it provides insight into why and when participation is theoretically and empirically connected to performance: it allows for better group decisions and inclusion into the group by allowing voice to those who do not agree. Participation viewed this way might be related to functional diversity as well, as

participation would allow various diverse viewpoints to benefit the work group as a whole. This is consistent with other considerations of participation in decision-making within organizations, such as Scott-Ladd, Travaglione, and Marshall (2006), who note an indirect positive relationship between participation and performance through increased affective commitment, job satisfaction, and performance effort.

  Interdependence. In groups, success and failure often occur at the collective level regardless of the actions of any one member, and the amount and type of interdependence of group members on each other is theorized by Campion et al. (1993) to be predictive of group outcomes. The CWGEM divides the interdependence theme into three sub-themes: *task interdependence* indicates that group members' performance of specific tasks depends on the performance of other group members, *goal interdependence* is the alignment of group member goals with collective group goals, and *interdependent feedback and rewards* occur when the incentive structure and performance feedback for a group member is informed by the collective performance of the group.

  Regarding interdependence, the initial study by Campion et al. (1993) found significant correlations of employee-based ratings of the interdependence theme and sub-themes with productivity. Neither the manager-based assessments of interdependence nor its sub-themes were significantly correlated with productivity outcomes. Employee-based assessments of *task interdependence* ($r = .14$; $p < .10$) and the composite interdependence theme ($r = .20$; $p < .05$) were both significant, indicating a small positive relationship between the level of inter-reliance group members have on each other and their overall productivity.

As previously discussed, Erez and Somech (1996), investigated performance levels of managers in different Israeli subcultures. Contrary to the negative findings of Campion et al. (1993), these researchers did find that clearly defined and challenging group goals based on collective performance of an entire team prevents loss of productivity within groups; *goal interdependence* was necessary for effective production (in this study, collective decision-making aligns individual goals with group goals), and collective group outcomes for success or failure resulted in *interdependent rewards and feedback*. Of note, Erez and Somech found that, regardless of other factors, group production is (as expected) higher with interdependent rewards and feedback than without it, and also is higher than the sum of production when all group members work as individuals. This means that groups can – under the correct conditions – be significantly more productive than the sum of their individual contributions.

Contrary to both the theory and results of Campion et al. (1993), De Drue and West (2001) found that neither *task interdependence* and *goal interdependence* had a significant relationship with desired work group outcomes (team innovation); it should be noted that these measures of interdependence were not the focus of De Drue and West, but rather factors that should be controlled for as initial correlates. De Drue and West, in this study into team innovation, provided evidence of the internal validity of these measures (Cronbach's alpha of .78 and .84, respectively), and note that team size, task interdependence, and goal interdependence together do not have a significant relationship with team innovation and account for a non-significant amount of variance in team innovation prediction. These findings are disputed by the results of Tarricone and Luca

(2002), who found that (consistent with the relationships theorized by Campion et al.) more interdependent groups functioned better than less interdependent groups.

Context. The effectiveness of work groups can be greatly influenced by factors outside the group; Campion et al. (1993)'s Work Group Effectiveness Model identifies three sub-themes in the category of context. These environmental or contextual factors include *training* (domain-specific training as well as training on how to work in groups, such as group process/interpersonal skills), *managerial support* (the parent or greater organization must support and invest resources into work groups for their success), and *cooperation and coordination between groups* (groups must be connected to the larger organization and interact with that organization as a distinct, identifiable group to be truly effective).

Campion et al. (1993) found only limited empirical support for these theorized relationships. Only the *managerial support* sub-theme correlated with increased workgroup production (with a small effect size). In terms of the productivity output of work groups, none of the subthemes (nor the context composite theme) were significantly correlated with employee assessments, and only the *managerial support* sub-theme (when assessed by managers) is correlated with increased workgroup production ($r = .16$; $p < .10$). It is possible that managers know better than employees how much support a work group is actually getting from managers (and then, by extension, one might infer that it is actual, tangible support from management that is important to performance in terms of this sub-theme, rather than feelings of managerial support a workgroup may or may not experience).

Campion et al. (1993) noted that organizational context is theoretically important but empirically under-investigated. Consistent with this claim, remarkably little generalizable empirical evidence exists that tests theoretical connections between this theme and work group performance outcomes. For a topic as oft-researched as training, studies seem to be unable to convincingly link it to group-level production when examining enduring, extant work teams. Those that do make the connection between training and performance in experimentally-assembled teams (such as Marks, Zaccaro, and Mathieu, 2000) have found success in showing relationships between performance and training *about group processes* (as opposed to training *about doing one's job*), which is the type of training recommended by Stevens and Campion (1994). In terms of *managerial support*, Miles and Mangold (2002) did not inspect the connection with group performance, but found that increasing managerial support increases employee satisfaction.

Composition. The characteristics of group membership can make a team more or less productive. The Campion work group effectiveness model (Campion et al. 1993) specifies that teams should have diversity in competencies, abilities, and backgrounds (especially when tasks assigned to the group are diverse) because the sub-theme *heterogeneity* increases the amount of knowledge transfer between group members. Additionally, team members should be able to accomplish multiple jobs within the group (*flexibility*), teams should be balanced in size to maximize potential worker output while minimizing managerial overhead or under-involvement of group members (*relative size*), and teams should ideally be comprised of members who have a *preference for group work*.

Campion et al. (1993)'s study found that employee's overall composite assessment of work group composition, as well as the sub-theme of *relative size*, were each positively correlated with work group production (modest correlations of $r = .21$ and .23, respectively; $p < .05$ for both). As assessed by managers, *relative size* was again positively correlated with productivity ($r = .19; p < .05$). These significant positive correlations between employee-assessed and manager-assessed measurements of relative size should not be interpreted to mean that the larger the group, the higher the productivity, but rather that more appropriately-sized groups are associated with higher productivity; the positive linear relationship is not between size and productivity, but instead appropriateness of size and productivity. One other significant correlation of note from the composition theme: employee assessments of *heterogeneity* were significantly negatively correlated with group production ($r = -.15; p < .10$), meaning that increased diversity within groups was associated with lower production. *Heterogeneity* of groups (and its related concept, diversity) remains complex and bears further discussion as to why it might sometimes be associated with decreased production, i.e. might some diversity be good for production while other types of diversity be detrimental?

Of the subthemes in the composition theme, most literature focuses on diversity (*heterogeneity*), often in the context of intragroup conflict, and sometimes tangentially addresses relative size. No research has been identified that explores group-level assessments of *preference for group work*; in fact Campion et al. (1996) recommends it be removed from the model. *Flexibility* is likely highly contextual and does not appear to be of interest to researchers, as evidence of its research has not been found.

Heterogeneity. Diversity within work groups in terms of member characteristics (e.g. *heterogeneity*, one of the subthemes of *composition*) is specifically researched because work teams can potentially be very homogeneous in certain aspects yet heterogeneous in others; researching what kinds of differences matter in a given context is critical for appropriate conclusions to be drawn. Pelled, Eisenhardt, and Xin (1999) tried to make distinctions between demographic diversity (differences between group members based on age, gender, ethnicity, etc.) and skill- or training-based diversity (work experience, functional expertise, organizational affiliation, etc.). Based on responses from 317 individuals from 45 teams belonging to the electronics divisions of three large corporations, Pelled, Eisenhardt, and Xin found that variations on most types of diversity (company tenure, age, ethnicity/race, and gender) were not significantly related to variations in group performance between work teams. However, these researchers found significant relationship between functional diversity and performance (which is to say, groups assembled from members with varied work specialties and skills performed better than more homogenous groups), though the effect of this relationship diminished as group longevity increased; this indicates that groups with members who bring different skills and employment experiences to a work group perform better initially (the study authors explain this through in terms of increased task conflict), but the differences tend to be less important over extended lengths of time. This accords with other studies such as Overbeck, Correll, and Park (2005), who provide evidence that diversity in job-related potential and ability is also important for efficient team functioning.

In a similar study, Jehn, Chadwick, and Thatcher (1997) also investigated the effects of various types of diversity on group performance through the explanatory

mechanism of intra-team conflict. These researchers studied 88 five-member teams of MBA students from three universities performing consulting projects. The types of diversity measured were demographic (age, gender, nationality, and education level) and value congruence (the level of similarity or difference between various personally-held values such as expectations of success); not measured were types of diversity such as group longevity (which was equal between all groups in this study) or functional diversity. This study did not find significant relationships between investigated measures of diversity and objective measures of group performance. This does not contradict Pelled, Eisenhardt, and Xin (1999) as their significant findings were not investigated by Jehn et al. Though Jehn et al.'s study found many significant correlations between measures of diversity and subjective or perceived performance, the study also found nearly no relationship between perceived performance and objective performance, meaning that actual objective performance and heterogeneity were not found to be significantly related.

In their previously-discussed investigations of diversity, the studies of Pelled, Eisenhardt, and Xin (1999) and Jehn, Chadwick, and Thatcher (1997) each investigated multiple types of intragroup conflict in relation to diversity and work group performance. O'Connor, Gruenfeld, and McGrath (1993) also investigated relationships between conflict and performance of work groups. O'Connor et al. reached different conclusions than Pelled et al. or Jehn et al. however, as they argued that increased levels of conflict does not usually increase work performance. They did not investigate differences between groups in terms of demographic heterogeneity but did note that conflict within any enduring group decreases temporarily when membership in the group changes.

Taken as a whole, these three studies provide evidence that while diversity of certain types is correlated with work group performance, evidence for the relationship between various types of conflict and performance is mixed and may be a side effect of other factors, such as diversity or duration of work group member stability.

Relative size. While the size of a work group is often considered in studies, it is usually to control for possible linear relationships between increasing group size and available work hours available to the group for production (e.g., De Drue & West, 2001), rather than an assessment of right- or wrong-sizedness. This is informative, because it may very well be that in studies of groups that vary in size, investigating whether the group is too large to be manageable or too small to be able to accomplish all requirements is important, but also important is the idea that there is an expectation that bigger groups can produce more than smaller groups due to increasing availability of worker-hours. Other studies such as Pelled, Eisenhardt, and Xin (1999), found no relationship between group size and performance.

Wheelan (2009) considers the problem of group size's relationship to performance through a model of group development, and argues that smaller teams can assemble and start being productive faster than larger teams. Wheelan uses a four-stage model that describes the coalescing and performance of new groups as they transition to performing groups; performance suffers in the first two phases but at stage three the group starts to perform, and may reach its potential in the final stage, stage four. Wheelan also connects the speed a group progresses through these developmental stages, explaining the negative relationship between them by noting that larger groups struggle to

coalesce as easily as smaller groups can (the author states that it should take an average of six months to achieve full performance at stage four for a new group).

Wheelan (2009)'s findings that smaller groups are more efficient seems limited. First, while possibly explanatory in terms of the rate a team may start to function at its potential, Wheelan's model does not address what that potential may be (in terms of group size). Wheelan does not include many other factors to control for (the author does check to see if there are differences between male-dominated and female-dominated groups, but does not control for any of the many known correlates that could have been included in a study on group performance). Perhaps Wheelan is best understood as considering a different type of group (Cohen and Bailey 1997's *project teams* instead of *work teams*) – but, it is unclear how Wheelan's model might be informative to extant, long-standing groups that endure periodic additions and departures of individual members. Finally, Wheelan makes a questionable assumption that the group's size is fixed at its formation. Wheelan's contribution to assessing the constituent parts of the Campion work group effectiveness model consists of considering the relationship between team size and performance, but does not address the relationship in a way that is compatible with or constructive to the discussion (too indirect and as a function of the rate of new team coalescence, not extant team performance).

Various scholars assess that work teams being overlarge (i.e. an excess in relative size) can be the precipitating factor to significant group dysfunction. Mueller (2011) argues that as team size increases, total production from the group tends to increase although individual contributions decrease, and assesses the primary cause to be loss of extrinsic motivation (in that, the larger the group, the more difficult it can be to use

extrinsic motivation or rewards on any specific individual) but finds little evidence to support a competing hypothesis that coordination loss is instead the primary cause (general inefficiencies generated by a larger – so more complex – system).

Process. The process theme is included in the Campion work group effectiveness model because work groups that function better internally will likely be more successful. Process sub-themes are *potency*, *social support*, *workload sharing*, and *communication and cooperation within group* (Campion et al., 1993). These groups need to believe they can be successful (potency), support each other (social support), ensure that all team members participate to an appropriate degree to prevent shirking or overloading (workload sharing), and have sufficient levels of communication and cooperation within the work group.

Of all the input themes, process is the most predictive of workgroup performance. All four subthemes (potency, social support, workload sharing, and communication/cooperation within the group) are positively correlated with performance when assessed by employees (with correlation coefficients ranging from .18 to .29; $p <$ .05 for each), as is the process theme composite measure ($r = .26$; $p < .05$). Managerial assessments were found to have similar relationships, with three subthemes (potency, workload sharing, and communication/cooperation within the group) having significant correlations with productivity ($r = .22$, .22, and .20, respectively; $p < .05$ for each). The managerially-assessed composite measure for process is also correlated with productivity ($r = .25$; $p < .05$).

While Campion et al. (1993) subdivides group-level processes into potency, social support, workload sharing, and communication / cooperation within group, other authors

– recognizing the importance of group-level processes – subcategorize differently. For example, Jordan, Field, and Armenakis (2002) explore group processes in terms of group potency, group cohesion, and team-member exchange. Group potency is common between the two, and it is not clear if there is much practical difference between Campion et al.'s other group process subthemes (social support, workload sharing, and group communication/cooperation) and Jordan et al.'s other group-level processes (group cohesion and team-member exchange); in fact an outside observer might conclude that they are just different ways to label the same intra-team dynamics. Jordan et al.'s study found that group potency, group cohesion, and team-member exchange to be highly inter-correlated (bivariate correlations ranging from .45 to .71; $p < .01$ for all three), which was consistent with Markova and Perry (2014), who found group potency to be highly correlated with group cohesion while investigating various predictors of cohesion and group wellbeing (but who did not include objective productivity as a criterion variable). Erez and Somech (1996) emphasize the critical importance of intragroup communication. Group potency was found by Jordan et al. to be more highly correlated with group performance outcome variables than was group cohesion or team-member exchange and was the most predictive variable of group performance, with a small effect size ($\Delta R^2 =$ .06; $p < .01$), after considering other group process variables, gender, and company tenure (group cohesion, in fact, did not explain any variance over and above potency and team-member exchange).

    Group potency. Group potency is a collective corollary of workplace self-efficacy. It is a jointly-held belief in the group's ability to succeed in general – at any task or challenge the group may face. Group potency is positively correlated with

collective success (Stajkovic, Lee & Nyberg, 2009) and has been demonstrated to be a reliable predictor of performance in semi-autonomous, self-managed teams (de Jong, de Ruyter & Wetzels, 2005). In a sales-specific context, Rego, Junior, and Cunha (2015) found that group potency explains sales performance over and above other potential explanatory variables (e.g., authentic leadership and group virtuousness).

Group potency can be measured at the group level, using focus groups and other communal assessment methods, or at the individual level, with scores usually aggregated as a group (Guzzo, et al., 1993). Individual assessments are likely preferable to group assessments for several reasons. Group assessments suffer from methodological concerns (false consensus reporting, etc.) and have less reliability than individual assessments. Jung and Sosik (2003) strongly favor individual assessments of group potency rather than group discussions, specifically noting that individual, private assessments of the group have greater predictive validity than do public group assessments. Guzzo et al. noted that when assessed at the individual level and aggregated, the variance of the group potency scores across all individuals in the group can be at least as informative as the mean score: groups that have a lack of consensus in their own potency have performance concerns that are at least as severe as groups that have low potency.

Collective efficacy. Related to, but distinct from, group potency is another construct called collective efficacy. The distinction bears discussion due to the common misunderstanding and confusion between the two constructs; they are often used incorrectly or interchangeably (Rego, Junior, and Cunha, 2015), or pooled together in a discussion of *group efficacy* (Jung and Sosik, 2003). Campion et al. (1993) noted that

efficacy is similar to potency. Chen and Bliese (2002) note that these two concepts are

highly related, and provide evidence (in the form of intraclass correlations) that groups

can be "differentiated reliably" based on collective efficacy.

Collective efficacy refers to the sense of ability and agency to accomplish specific

tasks when working through a team or collective social group (Bandura, 1997). This is a

distinctly different construct than group potency, because the non-specificity of task or

challenge is the key difference between the concept of group potency and the related but

distinct construct of collective efficacy (which is task- or domain-specific): group

potency assesses a group's joint-held belief in the likelihood of success in *general*, while

collective efficacy is a group's jointly held belief in the likelihood of success in a *specific*

task or functional domain (Stajkovic, Lee, & Nyberg, 2009). Both are predictive of

group outcomes: Stajkovic, Lee, & Nyberg note that measures of collective efficacy are

similarly relevant to collective performance outcomes as are measures of group potency;

Jung and Sosik (2003) conducted a two-wave longitudinal study and found that neither

construct was universally better at predictive group effectiveness, but noted that both

collective efficacy and group potency tend to grow over time as a group continues to

work together.   Collective efficacy is also related to individual efficacy: Fernandez-

Ballesteros, Diez-Nicolas, Caprara, Barbaranello, and Bandura (2002) conducted a study

in Spain and demonstrated that personal efficacy is a predictor of collective social

efficacy. The accepted standard measure for workplace collective efficacy is the

modified Jones scale as refined by Jex and Bliese (1999).

Workload sharing. The items used by Campion et al. (1993; 1996) are measures

of group members perception of the equity of workload distribution and an assessment of

group members' perceptions that group members as a whole were doing their fair share. Mulvey and Klein (1998) found that this perception of inadequate workload sharing (i.e. perceived social loafing) is extremely detrimental to group processes, decreasing performance through diminished motivation and commitment. Perceptions of social loafing are toxic: Comer (1995) argues that they increase the likelihood that others will loaf as well (because there is decreased incentive for individual commitment to the group while others are perceived to break group norms of commitment and effort). Mulvey and Klein note that perceived social loafing is universally detrimental and has no upside in team dynamics, lending credible support to Campion et al.'s inclusion of this subtheme in the CWGEM.

Summary of Initial Study Findings

Of the five input themes, Campion et al. (1993) found that job design and process were the most significantly correlated with objective work group production. Measuring input constructs from group member and from management often produces varied results; implying that there could be a disagreement between the two, or that one level is more appropriate than another for a particular theme or subtheme (though this is most likely highly contextual and might vary between organizations). Second, while four of the five themes' composite scores as assessed by employees were significantly correlated with performance outcomes, less than half of all subthemes were. One might infer that composite scores are more practically useful than all 19 subscales, or that certain subscales should be assessed specifically while many could be aggregated into a composite score without loss of explanatory or predictive power. Finally, in addition to developing a comprehensive theoretical model, Campion et al. also developed 3-item

30

measures for all 19 subscales; it is possible that some of the results (specifically the findings of non-significance) in this initial 1993 study are more reflective of the specific measures used. Non-significant findings in certain relationships might be interpreted as problems with the instrument rather than shortfalls in the theory.

Refinement and Extension in 1996 Replication Study

Campion, Papper, and Medsker (1996) replicated the 1993 study using a refined instrument and different organizational contexts. Instead of using similar groups from one company (a financial services company), this 1996 replication used 60 groups of "professionals" from three career fields (information technology, insurance, and administrators) across different companies (totaling 357 employees and 60 team managers). Two subscales (*task identity* and *preference for teamwork*) were removed from the model (reducing subthemes from 19 to 17) because the two subscales had little predictive relationship with group outcomes (Campion et al. 1993; Campion et al. 1996). This reduced the predictor variables to 5 themes and 17 subthemes. The dependent variables were two of the three outcome variables in the model: group member satisfaction and managerial evaluation and judgement of work group performance. Objective productivity metrics were removed as a dependent outcome variable from this study due to the difficulty of comparing tangible outcomes across organizational and functional boundaries. This study resulted in confirmatory support for the conceptual model (88% of the variance in the two types of work group outcomes was explained by the measured themes and subthemes in the work group effectiveness model).

Because the groups in this 1996 study produced different products and services, Campion et al. did not evaluate work group production. The authors concluded that work

groups with differing measures of objective productive performance were not readily evaluable against each other (i.e. while insurance teams can be objectively compared to each other, it is much harder to correctly compare objective productivity of an insurance team against an information technology service team), and they accepted this methodological issue as a limitation and did not consider objective production as outcomes. Because of this decision, this replication study cannot inform consideration of the relationships between the input themes and subthemes in the work group model and objective work group productivity. It can, however, provide evidence to support scale construction or correlations between predictors.

Accidental Convergence in 1997 Mixed-Methods Research

Many studies implicitly or explicitly investigate only a subset of the work group effectiveness model. No studies (beyond Campion et al. 1996) have been found that completely replicate Campion et al. 1993's model. Hyatt and Ruddy (1997) sought out to develop their own model using best practices from managers, which resulted in a model quite similar to the CWGEM.

Hyatt and Ruddy (1997) studied dispersed work teams (work groups physically spread out and conducting coordinated activities without co-location in a work space) using a mixed-methods design. While their methodology and theoretical starting point varied substantially from Campion et al. (1993), their results significantly converged, providing evidence of model validity for the Campion work group effectiveness model. These researchers (Hyatt and Ruddy) conducted in depth, qualitative interviews to identify likely antecedents of effective team outcomes and then conducted quantitative

correlational studies using hierarchical linear modeling and confirmatory factor analysis to model these relationships and provide evidence of scale validities.

The study investigated geographically-dispersed teams of service and repair persons specializing in computer and imaging machines (such as faxes and photocopiers). While this is an alternate theoretical model and a competitor to the work group effectiveness model of Campion et al. (1993), it is important to note that the qualitative interviews identified multiple themes advocated by Campion et al., and of all the characteristics of effective teams identified by Hyatt and Ruddy, nine of the 13 were attributed to Campion et al. or the precursor work of Hackman (1987) as a "similar model". Limitations of this study include a restriction of the work group outcome to managerial and customer satisfaction (i.e. no objective production was included as a criterion variable, such as the number of copiers fixed per month); this is similar to the limitation of Campion, et al. (1996). However, taken as a whole, this study appears to provide convergent evidence of the theoretical soundness of the work group effectiveness model.

Limitations

The limitations of the CWGEM are several: first, the entire model can be difficult and expensive to apply to a large organization; there are 17 subthemes (following the 1996 reduction from 19) to measure and it is not clear that measuring them all is better than choosing the most relevant to the organization and focusing on investigating a restricted model. This is clearly evident when considering the results of Campion et al. (1993)'s empirical study: using a 57-item, 3-item-per-subscale measure might have been too broad and too shallow an instrument. Perhaps theory and practical limitations of the

target organization should guide a more focused investigation of only some of these input themes and subthemes.

Second, though Campion et al. (1993)'s model allows for (but does not discuss or conduct) multivariate analysis of the relationship between predictor and criterion variables, it does not require or even recommend longitudinal assessment of any of its themes or subthemes, leaving it vulnerable to general criticism of the state of organization research levied by McGrath and Tschan (2007), resulting in a potentially distorted assessment of predictive relationships. McGrath and Tschan make several potentially worrying assessments of all non-longitudinal studies of organizations, such as the potential for lagged or leading relationships between predictors and outcomes (i.e. related concepts have to be measured at different times to understand their relationship) that cannot be detected without longitudinal studies, and the danger of using one slice in time as a representative sample of an enduring organization (i.e. non-longitudinal data should be assumed to have a sampling bias, under-representing data sets from all times not included in the study and over-representing available data from the one time included in the study). Additionally, there is some evidence that – at the individual level – contributions to one's group is dynamic and cyclical (Loch, Huberman, & Stout, 2000), reinforcing McGrath and Tschan's concerns.

Campion et al. (1993; 1996) are not immune to these criticisms about non-longitudinal studies, though there is no reason to view their research either as an outlier or somehow non-representative of top-tier organizational research. Additionally, longitudinal studies are rare, expensive, disruptive to the target organization and

potentially problematic if they cannot capture the change over time of all relevant variables, so there are compelling reasons to not perform them.

Third, while there are different types of teams (four types, according to Cohen and Bailey, 1997), the relationships described by Campion et al. (1993) appear to focus on the temporally stable, clearly defined *work team*. This limits the universal applicability of the model, but potentially allows for better predictive studies when examining types of teams and work groups that share these characteristics.

Finally, the wide and shallow research design that Campion et al. (1993) implemented involved testing bivariate correlations between each input (5 themes and 19 subthemes) and each output (3 types). As each input was measured two ways (group members and manager-supervisors), this results in 144 potential tests and the authors did not provide evidence of alpha level adjustment; repeated tests for significance likely caused some type I errors to occur. Also, individual correlations were small (ranging from .10 to .25), which correspond to small effect sizes. Considering these limitations, the strength of this Campion et al. study is that it demonstrates that one needs to look at organizations comprehensively and not search for one particular construct or measure that can radically change the outcome of a work group.

Social Value Orientation

In light of the reviewed literature, it is clear that the CWGEM is theoretically sound, generally predictive, but perhaps not universally agreed on regarding formulations of all constituent parts. Because of this, it may be worthwhile to consider a testable extension or modification of the model. One concept which presents extensive evidence

of predictive and explanatory validity to explain outcomes and behaviors is Social Value Orientation (SVO). As a theory, SVO explains interpersonal utility preferences.

Social Value Orientation is a personal trait that codifies an individual's preference of outcome balance between oneself and others in situations of social conflict; social conflict being any situation where different individuals may have different preferences for different outcomes. An individual's preferred combination of outcomes for others and themselves is called one's social value orientation. This discussion of SVO will argue for the inclusion of SVO as a potential axis of *heterogeneity* (i.e. groups may differ in terms of SVO, which is an important type of diversity) and certain subthemes of the process input themes (as SVO, at its core, is a trait that describes personal interactions in social settings).

Definition and Characteristics of Social Value Orientation

Social Value Orientation allows researchers to use game theory to assess a personality trait in individuals (Messick & McClintock, 1968) that has implications in the workplace (Upton, 2009). While research indicates SVO is a stable, lifelong personality trait that develops in children prior to adolescence (Knight & Dubro, 1984), researchers have routinely demonstrated the predictive power of SVO in various types of experimental games (McClintock, Messick, Kuhlman & Campos, 1973), relatively less work has been done in the workplace to advance development of practical applications of SVO. Some example studies include Bogaert, Boone & van Witteloostuijn (2012) who inspected the relationship between SVO, workplace climate, and affective commitment, and Upton (2009) who tested productivity of pairs of workers with matching SVOs and varying incentive plans. Lacking, however, is empirical or theoretical research on work

groups and collective measurements of SVO at a small-group level. Upton avoids this problem by using matching pairs of individuals, which is insufficient for real world, extant work teams.

Social Value Orientation is a relatively time-stable trait that develops as a child (Knight, 1981). However, the proportion of SVO in the population does shift gradually to be more pro-social throughout adulthood, with people valuing others more as they age (Van Lange, Otten, DeBruin & Joireman, 1997). One's own SVO influences one's ability to properly assess SVO in others through routine interaction (Maki & McClintock, 1983).

Measuring Social Value Orientation

Social Value Orientation is routinely assessed at the individual level using decomposed games. A decomposed game is created by collapsing a two (or more)-subject game (e.g. a Prisoner's Dilemma Game) into a one-subject choice that maintains the same payout matrix and tension, effectively "decomposing" the game into its constituent parts (Messick & McClintock, 1968). In short, by creating decomposed games, a researcher can use the advantages of game theory to ask survey respondents questions involving payout and outcome preferences without actually having to perform two-person experiments. By creating decomposed games, theoretical games are converted to survey items; SVO assessment instruments often use decomposed games. In situations where an actor can see an outcome in terms of results to self and to others, individuals have a preference as to the appropriate balance between the two – this preference is the defining characteristic of SVO and what makes it so compelling a possibility to include in a comprehensive model of work group performance.

Social Value Orientation is a categorical trait that assesses an individual's preferred outcome from a social conflict or social dilemma. The category represents the balance between benefit to oneself and benefit to others. The four orientations which people tend to have are a preference between are:

- joint outcome, reflecting a preference for mutual benefit and a desire to minimize the difference between one's own benefits and the benefits received by others;

- a preference for others' gain and an indifference for one's own gains or losses, reflecting an altruistic preference;

- a preference for one's own gain and an indifference for others, reflecting an individualist gain;

- a preference for one's own gain and other's loss, reflecting a competitive preference and a desire to maximize the difference between one's own benefits and those received by others.

The first two orientations (joint benefits and altruistic or others-only) can be grouped together and classified as *pro-social*, while the individualistic and competitive orientations can be grouped together and classified as *pro-self*.

Implications on Group Performance

Little research has looked at defining how to measure SVO at group level – as noted above – although Upton (2009) tested matching pairs working together (so the group level SVO was identical to that of both group members), while Bogaert, Boone & van Witteloostuijn (2012) structured their study to examine relationships between group-level measures of organizational climate and individual-level assessments of SVO, avoiding the challenge of aggregating SVO to the group. Group performance is likely

impacted by the SVOs of constituent members because group-level behavior is often predicted by individual measurements aggregated to the group level (as evidenced by the entire discussion of the Campion model and related literature, above). After considering Campion et al. (1993; 1996)'s model and related empirical studies, and the theoretical background and experimental performance of SVO, there are several compelling reasons to believe that SVO should be measured and integrated into a study which uses Campion's model as the conceptual framework.

In the workplace, people respond differently to production incentives based on SVO. In some situations, pro-selfs in groups will work harder, and in others, pro-socials in groups will work harder (Upton, 2009). This indicates that SVO might display an interaction with the feedback and incentive structure of the workgroup on the effect on productivity. Pro-socials will change their behavior to conserve collective resources, while pro-selfs will conserve resources only if it does not require substantial changes to behavior (Sutterlin, Brunner, and Siegrist, 2013). This indicates that groups with higher levels of pro-socials might be more efficient.

When groups compete with each other (intergroup conflict, is expected in many business and workplace situations; Mintzberg, 1983), pro-socials tend to give of themselves for the good of their group more than pro-selfs in the same group do (De Dreu, 2010). This indicates that groups with pro-socials might have a higher upper limit of productivity, but also might exhibit more signs of fatigue and burnout in the long run, potentially leading to lower long-term productivity. Pro-social people tend to cooperate with others unless they strongly believe that their cooperation will not be reciprocated, while pro-selfs are predisposed to expect that others might take advantage when possible

and not cooperate unless explicitly in one's self-interest (Van Lange and Leibrand, 1991). This indicates that groups with an even mix of pro-social and pro-self individuals might have more conflict and less aligned behavior than groups of either all pro-socials or all pro-selfs.

As a function of trust, pro-socials give more when they trust others, but pro-selfs tend to give less (Kanagaretnam, Mestelman, Nairar & Shehata, 2009). This indicates that groups comprised primarliy of pro-socials might function best when they display high levels of trust with each other, but groups that are majority pro-self might function best when they do not readily trust each other. In terms of group performance, this likely reinforces the inference that groups that are more evenly mixed between SVOs might perform different than groups that are more uniform.

Pro-selfs may be more likely to take actions that are harmful to others for their own gain in economic or workplace situations (called *economic opportunism* by Sakalaki & Sotiriou, 2012). These authors found a relationship between pro-self individuals and a predisposition to withhold information that would disadvantage themselves when applying for health insurance or selling a used car; pro-selfs also demonstrated an increased acceptance of going back on one's word in a business deal when a better offer presented itself. This might mean that groups with higher levels of pro-selfs trust their own group less and may not perform as well collectively.

Social Value Orientation is conceptually compatible with the CWGEM, in that measures of group members are explanatory of group behavior, and therefore group performance. For example, pro-socials are likely to have different levels of alignment to group goals than do pro-selfs, leading to varying levels of goal interdependence between

groups that are majority pro-social compared with groups that are majority pro-self.

Similarly, pro-socials and pro-selfs likely have different preferences for group work, as

the former are predisposed to be more group-centric than the latter. Differences in SVO

are a type of diversity that has not been researched before nor conceptualized as a

potential axis of *heterogeneity* within groups. Finally, while the process theme is the

most predictive part of Campion's model, subsequent research (Jordan, Field, &

Armenakis, 2002) found similar results using different subscales than Campion et al.

(1993).

As a result, it bears considering whether the entire process theme (excluding the

potency subtheme) could be replaced with SVO, which, based on the literature

considered previously, has the potential to explain differences in *social support*,

workload sharing, and *communication and cooperation within groups*, resulting in one

measure more efficiently replacing three. As a result, two testable propositions that can

contribute in answering research question one (is SVO predictive of work group

outcomes?) are:

- Proposition One: SVO should be expected to explain differences in work group
  productivity better than social support, workload sharing, and
  cooperation/communication within group (i.e. all process subthemes except for
  potency)

- Proposition Two: SVO should be expected to explain differences in work group
  productivity that is not currently explained in the Work Group Effectiveness
  Model

Model Revision

This section is primarily a consideration of potential changes or modifications to the model that will address the concerns and shortcomings identified in the first section. Much of this discussion is centered on Social Value Orientation and its potential as an alternative conceptualization of certain portions of the Campion Work Group Effectiveness Model. In addition to SVO, other changes following from relevant discussions in Section I are included. This third section concludes with the proposal of a revised model, which incorporates SVO and other changes into the model.

In addition to including a group-level measure of SVO as a predictor variable for work group performance, several other changes should be included based on the literature review in Section I (see figure 2):

- Job Design should be retained along with its subthemes, although task identity will be removed due to the lack of empirical support for its retention (and following recommendation of Campion et al. 1996);

- Interdependence should be retained, although the lack of much empirical support for subthemes in either Campion et al. study or other research leaves little reason to include subthemes;

- Composition should be substantially revised, as relative size is likely context-dependent yet predictive, flexibility is rarely a subject of interest for researchers, preference for group work was recommended for removal by Campion et al. (1996), and heterogeneity is a construct that is probably overbroad (different conceptualizations of diversity generate different empirical results) – as a result, composition as a theme will be removed, although, if appropriate to a study,

42

relative size or diversity would be excellent candidates for inclusion (although not
as two subscales to a consolidated scale such as "composition");

- Context is under-researched and there does not appear to be much empirical
  evidence to recommend retention in a revised model;

- Process is the most strongly-supported theme and should be retained, with
  potency remaining a subtheme and SVO being tested as a potential replacement
  for the other subthemes.

If this revision is successful, research question two (how can the CWGEM be improved?)
can be answered. One additional testable proposition follows from this question:

- Proposition Three: The revised model should predict differences in work group
  production between groups at least as well as the revised 1996 model.

Figure 2. Proposed Work Group Effectiveness Model Revision
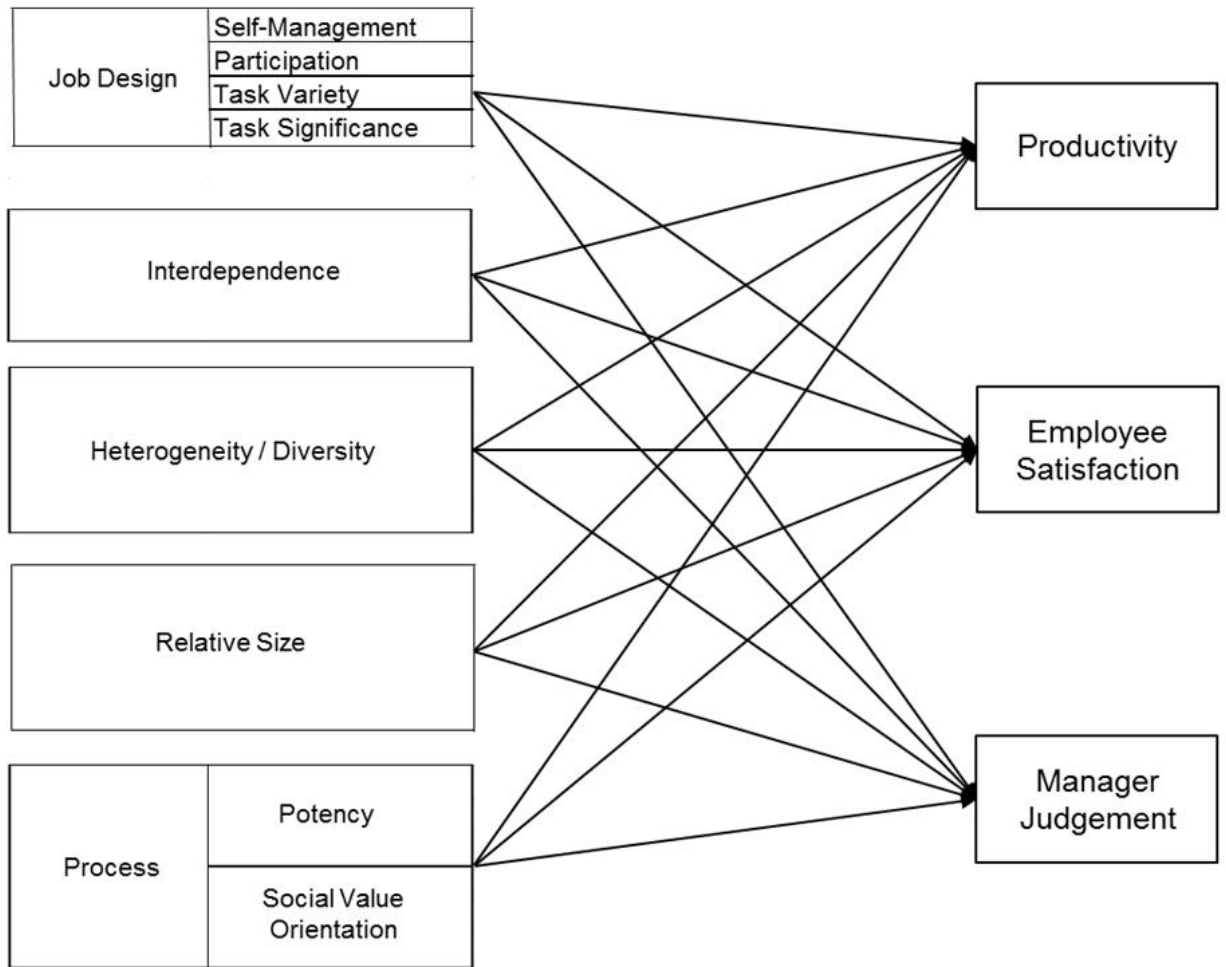
Figure 2. Depiction of revised relationships between inputs and outputs of work groups. The input scales and subscales are listed on the left side. The three outputs are listed on the right side.

As discussed in the introduction, the development of practical theories that can be used in real-world contexts is a desired outcome within the HRD community. The three outcomes of effective work groups and teams presented here are all important to organizations; productivity and positive manager judgements are business outcomes, while meta-analysis of employee satisfaction demonstrates substantial effects on reduced employee turnover, better customer satisfaction, and increased profitability (Harter,

Schmidt & Hayes, 2002). These outcomes are in fact reasons that organizations routinely establish work groups. Similarly, model input themes are functions of HRD decisions and practices – specifically recommending increased consideration in work group member selection and team structure, focused training for group members on effective group teamwork, and deliberate work process design. Work groups and teams are established by organizations to fulfill functions more effectively than multiple individuals could; better work groups and teams – the focus of this model – are important to any organization that uses work groups.

Specific implications of SVO's inclusion in practical HRD policies include a more deliberate consideration of personal traits of employees and their relation to work. Research into work group conflict (Kolb 2016) could benefit from an additional factor that explains interpersonal conflict (as SVO does). Future research into structural modeling and inspection of work group predictors as potential mediators or moderators is recommended. Creating and supporting work groups requires deliberate consideration and action; there is a clear business case for expecting managers and human resource specialists to use a model that connects inputs to desired outputs in order to maximize employee satisfaction and quantifiable outcomes.

Organizational Context

While the challenge of establishing and maintaining successful work groups and teams is common in many organizations, certain factors should facilitate the study of this problem. Ideally, one organization containing many hundreds of nearly-identical work groups and teams is desirable, because it would minimize extraneous factors and allow for the variables of interest to be more thoroughly explored. Additionally, the

organization's primary productivity should be a group-process output, and the organization should have identified the need to improve the performance of some or all of its groups. While limiting the study to one large organization potentially limits generalizability of any findings, attempting to use multiple similar organizations in the same study is methodologically troublesome. As an example, consider the difficulty experienced with Campion et al. (1996)'s attempt to include multiple organizations in the study while replicating Campion et al. (1993)'s study which only used one – key variables such as workgroup productivity were omitted.

Army Recruiting

One organization that meets the above criteria in terms of size, general homogeneity of work groups, and organizational output and need for improvement is the United States Army Recruiting Command. This organization (USAREC) is the part of the U.S. Army responsible for recruiting and enlisting new Soldiers into the Regular Army (full-time Soldiers) and Army Reserve (Federal part-time Soldiers); the National Guards (state militias) recruit for themselves and are one of USAREC's direct competitors (as are the recruiting activities of the other Armed Services). USAREC is subdivided into five recruiting brigades, each responsible for approximately 20% of the United States, with the coastal brigades also recruiting in overseas territories (such as Guam and Puerto Rico). The smallest organizational element in USAREC is the Recruiting Center.

Recruiting Centers are analogous to work groups and teams and meet the definitions discussed in Chapter I. In addition to sharing workspace and having assigned personnel and equipment, each center has explicit, unique monthly recruiting goals for

which all recruiters assigned to a center are collectively held responsible. There are approximately 1250 Recruiting Centers in USAREC and collectively they have geographic assignments built from all ZIP codes, each of which is uniquely assigned to one and only one Recruiting Center.

Recruiters in the U.S. Army are assigned to recruiting centers that are distributed geographically across the United States. These recruiting centers are each assigned ZIP codes that they are responsible for recruiting from; this assignment of land is both exclusive and exhaustive, which is to say, every zip code in the U.S. is assigned to one and only one recruiting center, with no overlap of ZIP codes across recruiting centers and no ZIP codes that are unassigned. Recruiters assigned to a given recruiting center are assigned a collective, annual recruiting mission they are expected to accomplish, often divided into monthly sub-goals.

Over the last several years, Army recruiting units have struggled to meet recruiting goals. One of the suspected causes of this is the decision in 2010 to change from individual recruiting (with individually-incentivized recruiting goals) to recruiting center-level recruiting (small team-based recruiting goals and incentives). This change from individual-level goals and rewards to group level goals and rewards predictably replaced poor cooperation between recruiters and unethical corner-cutting (hallmarks of group incentive structures) with lower overall productivity and lower aggregate per-recruiter participation and effort (hallmarks of group incentive structures; see Barnes, Hollenbeck, Jundt, De Rue, & Harmon, 2011). While the U.S. Army has a long history of successful team-based missions and activities, team-level sales force management remains a novel challenge. In an effort to increase efficiency of Army recruiting centers,

this study assesses the relationship between several measures known to predict group outcomes and the recruiting productivity rates of Army recruiting centers. The organizational benefit of this study is the identification of predictive factors indicating how and why some Army Recruiting Centers are much more productive than others, and potential development of effective interventions. In short, what recommended areas should USAREC invest time and resources in improving to maximize the benefits and minimize the detriments of the work group-centric recruiting operations due to the Army's decision to move from individual-level recruiting to small unit-level recruiting.

Context-specific predictors of productivity. In addition to the SVO and the variables of interest in the CWGEM, there are some additional variables that should be controlled for as they are either known or suspected to influence group productivity specific to either sales or Army Recruiting. This section of summarizes these factors and discusses their theoretical connection to productivity. These considerations primarily derive from organizational literature (technical reports and commissioned organizational studies) but are very much in accord with peer-reviewed sales force management literature such as Piercy, Cravens, & Morgan (1999), who discuss the practical effects that geography and population demographics have on the effectiveness of sales teams. The following discussion of these kinds of factors (and their implications to work group outcomes) is in the organizational context of military recruiting, though most of these concepts apply when considering sales team performance as well.

Moving from potential employees to actual employees (or potential recruits to actual soldiers) is called – by economists – conversion of potential supply. Rates of this conversion of potential recruits to enlistees are impacted by recruiting incentives,

48

quantity and quality of available recruiters, and managerial decisions and policies

resulting in the employment of these recruiters in more efficient or less efficient manners

(and factors that affect these rates must be controlled for). Orvis, Sastry, and McDonald

(1996) were very clear that efficient conversion of potential supply (i.e. what one might

call "effective recruiting") is as important to enlistment rates as the size of the population

or rates of propensity to enlist.

Segal, Bachman, Freedman-Doan (1999) demonstrated the relationship between

military pay, recruiting resources, youth unemployment, educational benefits provided by

the military with enlistment rates, indicating that conversion is higher when the military

is more competitive with other options. The performance of teams of recruiters is

influenced by the number of recruiters, quality of management overseeing recruiting

teams (Carroll, Lee, and Rao, 1986), population of the area assigned to the recruiting

team, and the balance between military pay and incentives compared to local youth

employment and academic opportunities.

Additionally, advertising and marketing is an important part of the military's

efforts to recruit. Advertising is both expensive and challenging to assess in terms of

specific return on investment rates, there is no doubt that its marginal costs are lower than

alternatives, such as additional recruiters (Dertouzos and Garber, 2003). Additional

research indicated that urban and suburban locations tend to produce more recruits than

do rural areas, while increased population density and per capita income are each

correlated with slightly decreased production (Mehay, Gue, & Hogan, 2000).

Based on this discussion, it is clear that while assessing performance of different

recruiting work groups, certain covariates should be included and controlled for: number

of recruiters in a recruiting center, total population of area assigned to recruiting center, and unemployment (specifically youth unemployment). Note that controlling for the number of recruiters in a group is different from the *relative size* subtheme of the *composition* theme in the Campion model; it is likely that there is a locally linear relationship between recruiters and production, but a certain number of recruiters is more effective on per-recruiter basis than centers that are too small to handle all requirements or too large for the center leader to effectively manage (as predicted by the *relative size* concept).

Propensity. The options available for potential Army recruits to choose from might be summarized as education, employment, or enlistment; propensity is the predisposition to make a choice to join the Army (and so, while this variable is binary categorical for an individual, when aggregated to a population it is a continuous measure of the estimated percentage of the youth population that are predisposed to join and would state so if asked in a survey). This choice can be influenced by many factors: Asch, Kilburn, and Klerman (1999) argued that the military is often in direct competition with universities for the best educated youth graduating high schools; Kleykamp (2006) conducted a study in Texas and found that expected increases in education costs often lead young people to use the military instrumentally to facilitate later education, further highlighting the relationship between the three options.

Orvis, Sastry, and McDonald (1996) defined propensity as an overall measure that summarizes an individual's initial interest to join. Propensity is a combination of (or perhaps the interaction between) key influencers, individual tastes, and the perceived youth labor market. Propensity can be general, relating to service in the military at large,

or it can be specific to a specific branch of the Armed Forces (Shavelson, Haggstrom, and

Winkler, 1983). Living near military installations increases propensity (Kleykamp,

2006), indicating that there are expected regional differences in propensity due to the

regional asymmetric distribution of military bases. Segal, Bachman, Freedman-Doan,

and O'Malley (1999) note that propensity is generally stable across time when measured

at the state or National level, though in some subgroups (based on ethnicity and gender),

generational shifts have been observed. Kleykamp also found that race/ethnicity rates in

the population were relative non-factors in varying enlistment rates, though lower

socioeconomic status (SES) does increase enlistment likelihood and initial propensity to

enlist.

The empirical research on the relationship between enlistment rates and

propensity is positively correlated and locally linear (locally linear meaning that the

relationship is approximately linear when inspected at current and likely future values,

but not necessarily linear across all possible values). The relationship is not fully linear,

however, due to the small amount of the target population that is propensed (currently

around 10%); the relationship between propensity and enlistment behavior displays an

elasticity effect. This elasticity can be quantified: Orvis, Sastry, and McDonald (1996)

note that a propensed young person is three times as likely to enlist in the Army as a non-

propensed individual, so changing propensity rates have different effects on enlistment

rates based on propensity, preventing a truly linear relationship (hence, elasticity). As an

example provided by Orvis, Sastry, and McDonald, decreasing propensity by 10% of

current levels would result in a decrease of about 2% of enlistments, while a higher

starting propensity would have a steeper decrease in enlistments given the same decrease in propensity.

CHAPTER III

METHODS

This chapter describes a study designed to answer the research questions proposed in the introduction chapter. The three testable propositions included at the conclusion of the literature review chapter are given an organizational context and a research study is designed around answering them. To review, the research questions and testable propositions for this study are:

- Research Question One: Is social value orientation predictive of work group outcomes?

- Research Question Two: How can the model used by Campion et al. (1996) be improved?

- Proposition One: SVO should be expected to explain differences in work group productivity better than social support, workload sharing, and cooperation/communication within group (i.e. all process subthemes except for potency)

- Proposition Two: SVO should be expected to explain differences in work group productivity that is not currently explained in the Work Group Effectiveness Model

- Proposition Three: The revised model should predict differences in work group production between groups at least as well as the original revised (1996) model

Participants

Study participants are Active Duty and Reserve recruiters assigned to USAREC Recruiting Centers. Every recruiter assigned to a Recruiting Center in the Third Recruiting Brigade will be asked to participate (i.e. census-style selection and recruitment of study participants). This constitutes 1704 recruiters assigned to a total of 275 Recruiting Centers. While it might be preferable to randomly sample 20% of the Recruiting Centers in all five brigades, organizational restrictions limit the study to only the Third Recruiting Brigade.

The geographic area of responsibility assigned to the Third Recruiting Brigade is the Midwestern United States, ranging from the Dakotas to the Great Lakes to Tennessee. This area has the full range of locales found across the U.S., from dense metropolitan cities such as Chicago and Detroit to rural and frontier regions such as Michigan's Upper Peninsula.

Figure 3. Map of Third Recruiting Brigade, United States Army Recruiting Command
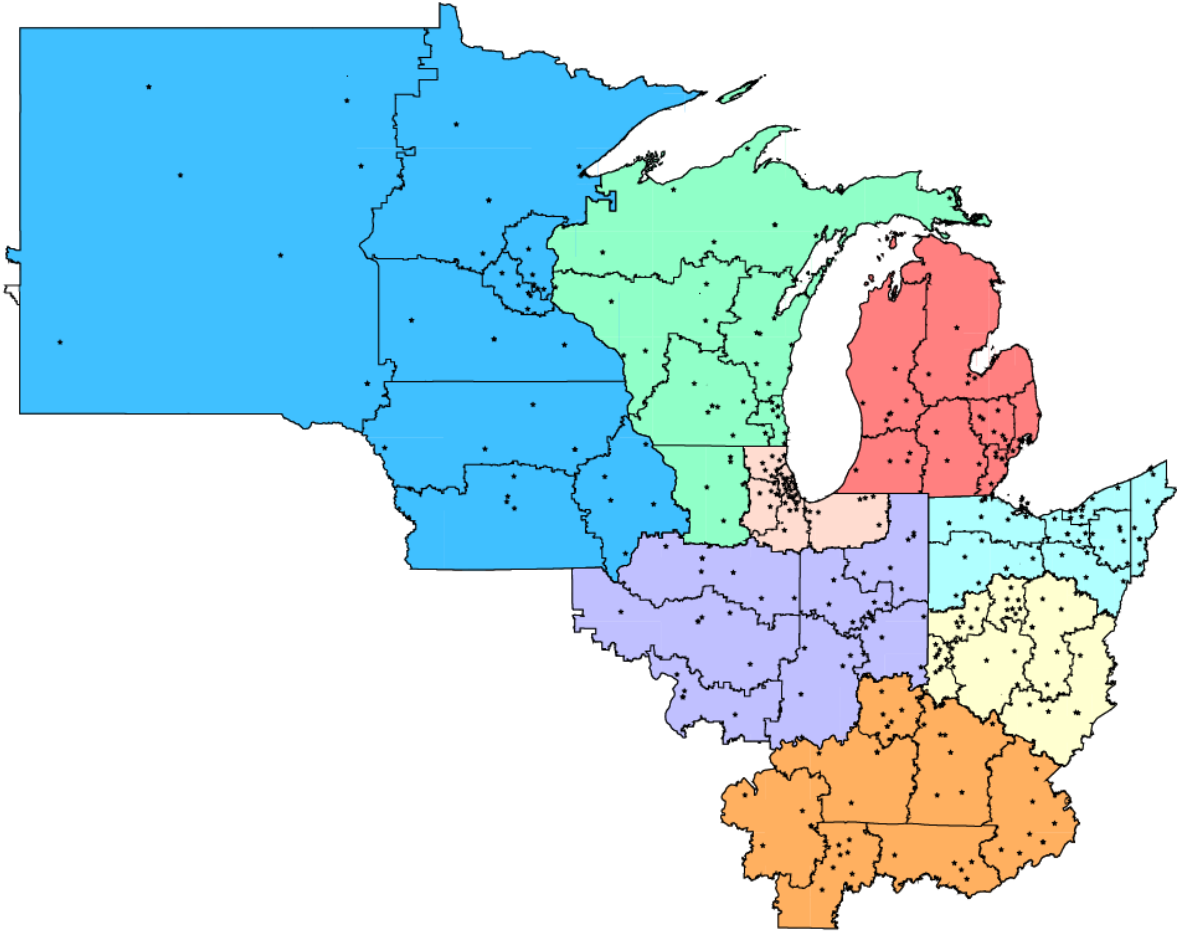
Figure 3. Map showing the physical locations of the sample in this study. The entire depicted area (from the Dakotas in the upper left of the figure, the majority of Kentucky and Tennessee in the lower right, and northwest Pennsylvania in the center-right of the figure) shows the eight recruiting battalions (various colors), the 54 recruiting companies (bounded by black lines) and the 275 recruiting centers (black dots). Courtesy of Headquarters, Third Recruiting Brigade, Fort Knox, KY.

Participant characteristics. Army Recruiters are career enlisted Soldiers that initially are temporarily assigned this duty away from their military specialty for three years, many of whom then voluntarily change their career specialization to become recruiters permanently for the remainder of their careers. Those that do not, and are on

this initial temporary duty, are called *detailed recruiters*. Most are on Active Duty although some are in the Army Reserve. Army-wide, the roughly 10,000 recruiters are primarily male (the Army as a whole is about 13% female; United States Department of Defense, 2012). The Army's enlisted force (of which all these recruiters are a part of) is 22% Black or African American, 3% Asian, 1% Pacific Islander, 1% Native American. The recruiters in this study are between the E-5 and E-7 pay grades. Because the percentage of minorities increases with seniority, E-5s and E-6s (together) are approximately 34% minority while E-7s and above are 47% minority. These recruiters are professional, career Soldiers. About 80% are serving on Active Duty; the other 20% are Army Reservists. Their ages range from around 25 at the youngest through late 40s for some of the older Army Reservists.

      Sampling procedures. A blanket (or census-style) survey contacted all Third Brigade recruiters via email with a request to complete a survey. These recruiters were solicited for responses through their official government-provided email address, which they use on a daily basis - in this case, defining the population as all recruiters in the Brigade, all members of the population will be contacted and solicited. If the population is considered all recruiters in the Army, however, the 1704-Soldier sample could be considered as a convenience sample – subject to organizational restrictions on survey access. Both of these population considerations is important, as the Third Recruiting Brigade is more likely than others to implement findings due to membership and influence in the organization, although, if successful, other Brigades will want to be able to generalize our findings to their organizations.

Potential error inherent in population. There are some unique characteristics of this population, but these generally do not appear to greatly increase the risk of a coverage error. All recruiters are English-speaking, literate, adult American citizens with no Service-limiting physical handicap. Even so, Army Soldiers are likely over-surveyed and have survey fatigue, which might lead to lower response rates and lower effort in accurate and thoughtful responses. This survey will address this with a personal appeal from Brigade leadership (in this case, from the senior leader in the Brigade – the Brigade Commander), and an explanation of how this research will directly impact policies which directly affect recruiters: these are both explicit in the pre-notification memorandum and the cover letter. Each recruiter has government-provided electronic devices, such as a laptop and a computing tablet (and has been trained to a base level of proficiency on them), so a web-based survey can reach them easily without intermediate gatekeepers or requiring additional training.

Mode of Administration

The primary mode of administration for this survey is web-based (online and self-administered). Primary reasons for this decision include the geographical dispersion of Army recruiters and the desire to acquire contemporaneous data (i.e. a "slice in time" across the whole organization) where spending weeks or months travelling around to the 275 Recruiting centers and administering the test in person would result in significant time delays between the first and last surveys. This might lead to construct-irrelevant variance as recruiters might talk about the survey to each other, a possibility that increases the longer the survey is open.

The headquarters at USAREC hosts a server running Electronic Feedback Management (Verint, 2014) survey software which was used to distribute the survey. This server emails the survey with unique access URLs to each study participant's government cellphone and laptop. To increase survey participation, the email sender the software attaches to the survey is routinely a senior leader (i.e. USAREC commander or Brigade commander); because this survey is only going to the Third Recruiting Brigade, the sender will be the commander of the Third Recruiting Brigade.

This survey methods should ensure excellent coverage. Due to the Army's hierarchical nature and routine communication requirements, response rates of over 60 percent are generally expected. Conversations with personnel who distributed surveys to the Third Recruiting Brigade using SurveyMonkey software in 2009-2011 revealed that response rates of 60 to 80 percent were common (Dr. Michael Benver, personal communication, August 2016).

Variables

Each variable will include a both a conceptual discussion and an operational description. Each variable will be identified as a dependent variable, an independent variable (of interest), or an independent variable (to be controlled for; a covariate). Finally, the source of the variable will be noted. Variables in this study include information from three sources: survey results, USAREC production databases, and other demographic information. Unless otherwise noted in the variable type subheading, all variables are measured and considered at the recruiting center level of analysis. Figure 3, below, clarifies these relationships.

Figure 3. Contextual and General Predictors of Work Group Effectiveness.
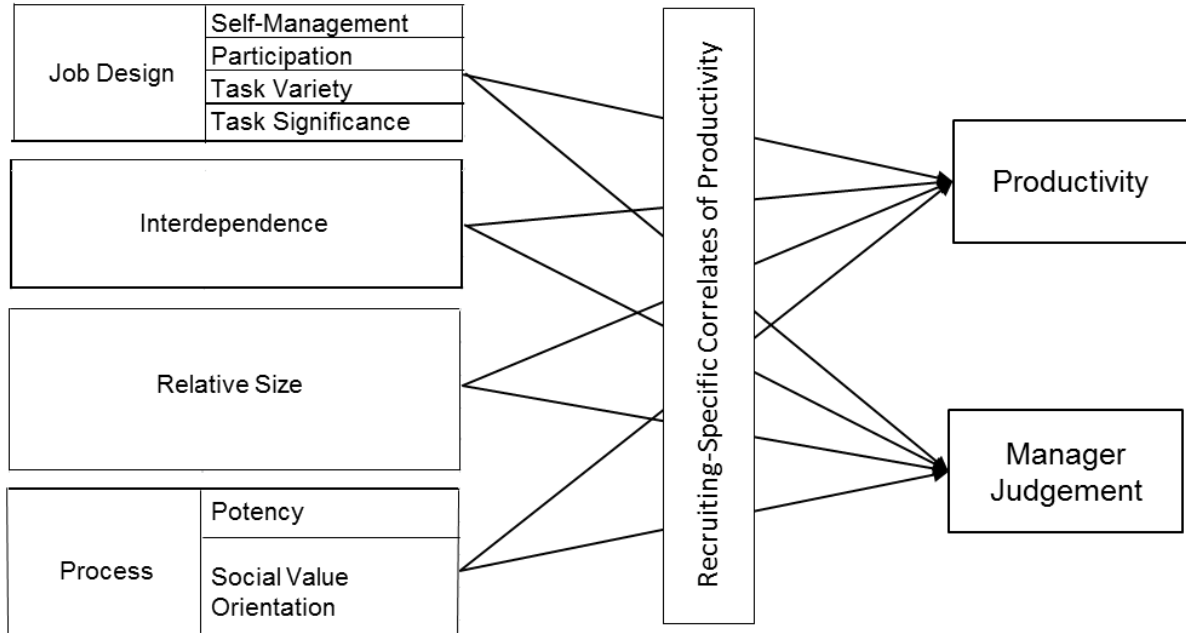
Figure 3. Depiction of theorized relationships between inputs and outputs of work groups, noting that all inputs can only be understood in the context of organization-specific predictors which will be controlled for.

Recruiting production. This is the fundamental, quantifiable output of Recruiting Centers. In the context of recruiting operations, this is conceptually the same objective production that nearly all of the work group studies reviewed in chapter two investigated in other contexts. Their production, aggregated from all 1250+ Recruiting Centers each month, is the total new manpower enlisted into the U.S. Army.

Enlistments come in two main categories, Regular Army and Army Reserve. These two categories are further subdivided into – for Regular Army – graduate alpha (high school graduates in the median top 50% of scores on the Armed Services Vocational Aptitude Battery, or ASVAB), senior alpha (seniors scoring in the top half of the ASVAB), and other (all other applicants, either low-ASVAB scores, GED-holders, etc.), and – for Army Reserve – graduate alpha, senior alpha, prior service (former

Regular Army Soldiers interested in the Army Reserve), and other. Collectively, these seven missions can be summed up into a *volume* mission. The amount of recruits enlisted is referred to as *production*. The volumetric number of recruits across Army components (Regular Army and Army Reserve) is the total, combined production number that is considered in this study; this production is always associated with a measure of time (e.g. fiscal year-to-date production, production per month, etc.).

Operational description. A common measure for a recruiting team's production for this is total recruits per month or year (for example, if a recruiting center recruited a total of 180 people into the Army in Fiscal Year 14, we could say that the center's annual production was 180 and average monthly production was 15). A rule of thumb for a "good" production value for this is around one recruit per recruiter per month – though most recruiting centers have routinely fallen far short of this since changing from individual to team recruiting goals (current averages across the country are around .8 recruits per recruiter per month).

For this study, the first three quarters of Fiscal Year 2017 (October, 2016 through June, 2017) is used. The sum total production for each center is scaled to a monthly, per-recruiter rate by dividing the total production in these nine months by nine and again by the number of recruiters each center is designed to have. As an example, if a three-person center had a productivity rate of .67, this should be interpreted as this center has produced, on average, 2 recruits per month for the first nine months of Fiscal Year 2017.

Variable type. This is one of two dependent variables of the study. Factors that are correlated with higher or lower recruiting production are of primary concern.

Source. This production data is routinely captured and stored via computer as part of the recruiting process; unlike SVO and group potency, both of which require application of an instrument to assess, recruiting production data (from 2010 to present) is readily available through database access queries, and historical data prior to 2010 are archived on paper-based reports.

Mission accomplishment. The second way to assess the tangible performance of a work group is to compare it to a standard set by management (Campion et al. 1993's *Manager Judgement* variable). Missions (for months and for years) are entered into a computer tracking program that all members of the organization have, so that everyone knows that their recruiting goals are well in advance. Production is tracked alongside of the mission, so that where each unit stands is known to leaders at all levels of the organization.

Operational description. This is the total mission accomplishment for the first nine months of Fiscal Year 2017; the total number of recruits a center has produced divided by the total number of recruits they were told they were expected to produce.

Variable type. This is one of two dependent variables of the study. Factors that are correlated with higher or lower mission accomplishment are of primary concern.

Source. Tracked by the USAREC recruiting database and applications.

Group potency. Group potency is a collective assessment of a group's ability to succeed (Guzzo, Yost, Campbell, & Shea, 1993). Group potency is the collective corollary of workplace self-efficacy. It is a jointly-held belief in the group's ability to succeed in general – at any task or challenge the group may face. The non-specificity of task or challenge is the key difference between the concept of group potency and the

related but distinct construct of collective efficacy: group potency assesses a group's joint-held belief in the likelihood of success in general, while collective efficacy is a group's jointly held belief in the likelihood of success in a specific task or functional domain (Stajkovic, Lee, & Nyberg, 2009). Group potency is positively correlated with collective success (Stajkovic, et al.) and has been demonstrated to be a reliable predictor of performance in semi-autonomous, self-managed teams (de Jong, de Ruyter, & Wetzels, 2005), which are fairly analogous to the structure of Army recruiting in terms of geographically dispersed and managed recruiting centers.

Group potency can be measured at the group level, using focus groups and other communal assessment methods, or at the individual level, with scores usually aggregated as a group (Guzzo, et al., 1993). Individual assessments are likely preferable to group assessments for two reasons. Group assessments suffer from methodological concerns (false consensus reporting, etc.) and have less reliability than individual assessments. Also, Guzzo et al. noted that when assessed at the individual level and aggregated, the variance of the group potency scores across all individuals in the group can be at least as informative as the mean score: groups that have a lack of consensus in their own potency have performance concerns that are at least as severe as groups that have low potency.

Operational description. The instrument for this variable is the eight-item group potency scale found in Guzzo, et al. (1993). This scale is specifically designed to measure the jointly-held beliefs about the group's ability to succeed in any task or overcome any challenge. It also measures the degree of intra-group alignment, and a high level of disagreement within a work group in this measure (as indicated by insufficient inter-rater reliability scores, r_{wg}), will indicate a common or jointly-held

opinion may be lacking. This can provide crucial insight into a dysfunctional team (Stajkovic, Lee, & Nyberg, 2009).

The normative sampling and results of the group potency measure provides evidence of validity: the internal consistency of the items was reported as .88 (Guzzo et al., 1993). The normative groups had high alignment within teams ($r=+.95$) indicating common opinions within groups. The normative groups also were distinguishable from each other (Guzzo, et al.). Other researchers have used this measure successfully, including de Jong, de Ruyter, and Wetzels (2005), who provide evidence of convergent reliability through ten face-to-face, in-depth interviews to ensure that the qualitative statements made converged with the results of the items on the scale.

This group potency measure is an eight item, closed-ended instrument; all items are of the same format and use a five-point Likert-type scale. One sample item is: "As a team, my recruiting center has confidence in itself." The scores are averaged and a mean and variation is calculated easily. The precise wording of the eight-item Guzzo et al. (1993) scale may include verbiage that is unclear to Army recruiters. For this reason, minor modifications to terms will be made. The term "this team" is expanded to indicate "as a team, my recruiting center…" or "…my recruiting center, as a team." The last two words on item eight ("around here") may cause ambiguity hence it will be omitted, or changed to "in my Recruiting Battalion" or "in the communities where we recruit." The instrument, located in Appendix A, notes all minor word changes (for this and for other scales).

Variable type. This is an independent variable of interest. Higher levels of group potency are expected to be associated with higher levels of recruiting productivity.

Source. Survey responses.

Social value orientation (SVO). SVO is personality trait characterized by outcome preferences in situations of social conflict (Messick & McClintock, 1968). In situations where an actor can see an outcome in terms of results to self and to others, individuals have a preference as to the appropriate balance between the two. This can be conceptualized as a two-axis graph, with one's own outcome on the x axis and the other's outcome on the y axis, and positive and negative signs indicating desired or undesired utility (Liebrand & McClintock, 1988; Knight & Dubro, 1984a). Valuing other's outcomes positively and demonstrating an indifference to one's own outcomes is an example of altruism, for instance. The opposite, valuing a negative outcome for others while maintaining indifference to one's own outcomes might be called aggression. Competitors seek to maximize the difference between themselves and others, so value positive utility for themselves and negative utility for others. The other five combinations are similarly labeled (see Maki & McClintock, 1983 for a chart of Griesinger & Livingston, 1973's original formulation of these axes). In practice, however, only four of the eight orientations are found to exist: altruists (positive to others, indifferent to oneself), cooperators (positive to others, positive to oneself), individualists (indifferent to others, positive to oneself), and, rarely, competitors (negative to others, positive to oneself; Liebrand 1984).

Operational description. The Triple-Dominance Measure (Van Lange, et al., 1997) will be used to measure SVO of study participants. The Triple-Dominance Measure's purpose is to assess the respondent's SVO preference (though not strength of preference; Murphy & Ackerman, 2014). This measure is unidimensional and does not

include subscales. The nine items are all structured identically: three options per item and all items are closed-ended, multiple-choice type (select one), and each option on each item includes payout pairs of varying values. The payouts are labeled for oneself and for one's partner. The three options for each question vary from each other in the following manner: one option has the highest total payout sum (i.e. payout for self plus payout for other is maximized), one has the highest individual payout sum (i.e. payout for self is maximized), and one has the highest difference between individual and other payouts (i.e. payout for self-minus payout for other-is maximized). The entire measure and prompts for this measure are included in Appendix A.

The scale for assessment is cutoff-based: if six or more of the nine answers correspond to the total payout sum, then the respondent is labeled *pro-social*, if six or more correspond to the total self-payout, then the respondent is labeled *individualist*, and if six or more correspond to the maximum difference, then respondent it labeled *competitor*. If none of the three categories have a score of six or more, the respondent it labeled *inconsistent* and is normally excluded from future testing. An important caveat: the label is the result of this scale, and the total score is not recorded (i.e. six of nine for pro-social is recorded the same as nine of nine for pro-social, i.e. just *pro-social*). This measure provides a single, categorical outcome.

Recent studies have successfully used the Triple-Dominance Measure and have reported satisfaction with the psychometrics. Upton (2009) administered the Triple-Dominance Measure to 182 undergraduate business students in the U.S. and reported that only 18 (less than 10%) were classified as "inconsistent", meaning that the measure was internally consistent enough to evaluate the SVO on over 90% of the time. Other recent

research reinforces this evidence of internal consistency: in a small, 18-subject study, van den Bos et al. (2009) reported that all medical subjects were consistent, scoring six or more of the nine-item Triple-Dominance Measure; Bogaert, Boone, and van Witteloostuijn (2011) reported that 85% of their study (n=209) received a classification that that this assessment of SVO produces a test-retest stability rating of 60-75 percent (indicating that individuals assessing their SVO at more than one time in their lives score the same more often than not); Sakalaki and Sotiriou (2012) used the Triple-Dominance Measure translated into Greek, and reported that 78.6% of 125 Greek undergraduates were successfully classified; Yamagishi et al. (2013) utilized the Triple-Dominance measure in several experiments, consistently reporting 95% classification rates. Taken together, this is strong evidence that the Triple-Dominance Measure accurately measures SVO. Triple-Dominance Measure of SVO (Van Lange, et al., 1997) offers the best balance of reliability, scholarly rigor, and ease of use (compared to competing measures of SVO – see Murphy & Ackermann, 2014, for a detailed assessment of all extant SVO scales).

Variable type. This is an independent variable of interest. Because Upton (2009) found that the relationship between the majority SVO of a group (he used 2-person groups of matching SVOs) and workgroup productivity is contingent on other work group factors (such as the type of incentives in Upton's study), I expect that SVO will be related to work group productivity although the direction of the relationship is unclear.

Source. Survey responses.

Propensity. This is the stated predisposition of young adults to consider serving in the Armed Forces.

Operational description. This is collected twice a year from high school seniors as part of youth attitudinal polling done by Department of Defense researchers. One four-point Likert-type item is used, and the two responses that indicate unwillingness to serve (strongly or mildly) are pooled together as non-propensed, while the two that indicate some stated willingness to serve are pooled together as propensed. Estimates based on demographics and geographic location of the respondent are used to weight responses and create estimates. The propensity for a geographic region is the average of the last two survey waves (waves occur two times per year). At the individual level, propensity is a binary categorical assessment; at the aggregate level it is a continuous measure of the estimated percentage of the youth population that would have been assessed as propensed had they been surveyed.

Variable type. This is an independent variable that is used at the recruiting battalion level. Due to small sample sizes below the state- or recruiting battalion-level, the measurement error in propensity becomes too large when considering smaller areas.

Source. USAREC data warehouse.

Unemployment rate. This is the percentage of people who want to be working that are not. Individuals that are not seeking employment are not counted as unemployed.

Operational description. This is the U.S. Government's Bureau of Labor Statistics (BLS) U-3 unemployment number. It is an estimate of current unemployment for all adults in the US. A better operationalization for recruiting purposes would be the expected (rather than current) unemployment rate of young adults (ages 17-24) rather than all adults. Since this expected future youth unemployment rate is not readily available, current unemployment rates for all adults are used.

Variable type. Unemployment is an independent variable and used as a control variable. This variable is not available at the recruiting center level. The lowest level of availability is at the recruiting company level.

Source. Ultimately from the BLS, it is calculated and estimated by USAREC's data warehouse.

Population. The number of people available to potentially recruit in a given geographical area.

Operational description. This is an estimate (often called QMA, or Qualified Military Applicants) of the young adult (17-24 or 17-29) population. Using the 2010 census as a starting point, and estimating population shifts each year since 2010, QMA does not include non-eligible persons in the target age range (such as persons already serving in the Armed Forces or those who are incarcerated).

Variable type. Population is an independent variable and used as a control variable.

Source. USAREC's data warehouse.

Number of recruiters. This is the number of recruiter positions allocated to specific recruiting centers.

Operational description. The number of recruiters is the designated number of recruiters that each individual recruiting center is specified have (called *Required Recruiting Force*, or *RRF*). This is not an actual count of people, as a center might be shorthanded, over-strength, have recruiters on leave, in schools, sick, suspended from recruiting duty due to misconduct, or otherwise unavailable to perform recruiting functions as a contributing member of a team. There are advantages to using this

number, however, as all other resources and allocations for a center (such as physical size of center, number of vehicles, and some advertising resources) are a function of this required number of recruiters. This number is also fairly stable, as it only changes when resources or territory are realigned through a formal, deliberative process.

The second way to operationalize the number of recruiters is to use the *foxhole strength*, which is a guess of the actual number of recruiters available in a given month. The number of recruiters who are unavailable to recruit for at least half of a given month, is subtracted from the number of recruiters assigned to a center, resulting in an estimate of the number of recruiters actually recruiting. This is not as superior a method of counting recruiters as it initially may seem, however: because this is a monthly estimate, and anyone available from 51% to 100% of a month is counted as available for the whole month, some error is introduced that is not necessarily less than the error between units from their required number of recruiters and the actual number they have. While this second way to count recruiters (available instead of required) is an attractive option, it is a trailing number (i.e. assessed after the fact, unlike required recruiter strength, which changes through a slow and deliberate process, and changes are known in advance). For these reasons, the required number will be used instead of the available number.

Variable type. Independent variable will be used as a control variable. Likely highly correlated with population (as population one of several factors in USAREC's model that distributes recruiters between recruiting centers).

Source. Required recruiting numbers are documented in each centers' documented authorizations for resources.

Variables from the CWGEM. Variables that Campion et al. (1993; 1996) used that are retained in the current model will use the scales found in Campion et al. (1996). Group potency will use the complete Guzzo et al. (1993) scale – as discussed above – so is not included in this current discussion. However, job design, interdependence, relative size, social support, workload sharing, and cooperation and communication within the group are all included in this study as independent variables.

Operational description. The measures used by Campion et al. (1993; 1996) will be used for these variables

Variable type. Independent variables.

Source. Survey.

Research Design

This quantitative, correlational study uses pre-existing work groups. Naturally-occurring variances in predictor variables will be correlated with naturally-occurring variances in the criterion variable. The study is a between-groups design, with the groups being organizationally-defined preexisting work teams (Army Recruiting Centers). Recruiters were assigned by the organization to their current Recruiting Center effectively at random in regards to SVO or other considerations in this study, but were not assigned by the researcher. Because this is a correlational study, there is no assignment of study participants to treatment groups or control groups, any interventions, or any experimental manipulations; a follow-on study could use interventions and treatment / control groups, however, if the correlational study demonstrates a relationship and the organizational decides to allow manipulation of recruiter assignment for organizational effectiveness and research purposes.

Procedures

  Data collection. Survey data is collected initially by survey software. Organizational performance and population data is downloaded from the organization's data warehouse. These data sets are merged at the recruiting center level and cleaned in Excel before export via CSV for use in R.   Before use, all data is inspected for appropriate ranges and responses using descriptive statistics.

  Data management. Data is stored on government-issued computer, and backed up on an off-site virtual drive. All identifying information is removed from individual responses and random key identifiers used for each individual response. All information is aggregated to the work-group level.

  Contact plan. The survey is administered via electronic mail: accompanying the survey welcome/introduction letter is a unique survey link. The survey is open for roughly two weeks with weekly reminder notifications sent to those who have not yet completed the survey.

  Threats to validity. One threat to validity is the fact that recruiters are very busy and may not put much effort into completing the survey accurately and with consideration. They may click-through the web-based survey haphazardly to just get it done. To combat this, all messaging and communication about the survey emphasizes how directly relevant it is to policies that higher headquarters develop; completing this survey accurately should provide information for organizational leaders to make data-informed decisions improving efficiency. Also, the sender field of the email is the Brigade Commander, taking advantage of the Army's hierarchical nature to encourage thoughtful participation.

Validity may also be compromised if recruiters talk to each other and try to game the answers by answering a certain way. If this happens, then the instruments would not measure individual recruiter-level differences the way they should. This is unlikely, however, as the survey is a conglomeration of multiple instruments and the recruiters get surveyed often enough that filling out a new survey should not be novel enough to initiate rampant collaboration or cooperation. The research results should be generalizable to Army recruiters worldwide; generalizability to other Department of Defense recruiting activities or other work groups and teams in governmental, for-profit, or non-profit settings will be more limited and accompanied with caveats.

Methodological concerns. There are four different methodological issues to consider when studying work groups in terms of inputs and outputs. Aggregation of individual scores to group scores must be done with deliberation and clearly explained. Group measures that are assessed from multiple members should be treated in a fashion similar to calibrating multiple observers of an object of study; correlation should be expected. Most advanced statistical tools (multivariate analysis, factor analysis, and interaction effects) require many more degrees of freedom than do simple bivariate correlations; there is a cost to the organization in terms of time and disruption and often real-world experiments are limited in size and scope below what is required for desired statistical analysis. Finally, the common practice of assessing work group antecedents and outcomes through quantitative measures may be limited, especially when drawing comparisons across organizational boundaries. A more detailed discussion of each of these four points follows.

72

Measures of individuals within groups are problematic because they must be aggregated to a group score to be considered alongside group-level variables. One example is Hecht et al.'s (2002) consideration of how to consider college students' grades on a group project with individual grades they received in the course prior to the experiment; these researchers chose to take the mean, although other options are using the group supervisor's individual score as a separate variable from the average of the remaining group members' individual scores (LePine et al., 1997) or taken the highest or lowest individual score and used that as the group score (Barrick et al., 1998). Each construct and measure should be considered individually, and aggregation to the group level should be conducted in a way that is thoughtful and theoretically plausible; the decision should be discussed and justified as part of the research study.

When individual group members are all scored on their assessment of a group value or process, such as group potency, all group members are individually rating the same value, so should be held to the same methodological expectations as other research settings when multiple raters or evaluators are employed to assess the same value. One way to assess the level to which multiple group members agree is with the inter-rater reliability coefficient (annotated as the within-group reliability coefficient, or r_{wg}) (Castro, 2002; James, Demaree & Wolf, 1984; James, Demaree & Wolf, 1993). When individual members are assessed on a group-level construct, a lack of correlation as between the members (indicated by a small r_{wg}) indicates that the measure may be unsuitable to use as a predictor variable in the study and is evidence against reliability. This study uses Lindell, Brandt, and Whitney (1999)'s modified scale $r_{wg}*$, which is

identical to r_{wg} for single-item scales but updated for multi-item scales; Newman and Sin (2009) conclude that neither r_{wg} nor $r_{wg}*$ is recommended over the other in all situations.

While some work group research is conducted in experimental settings (usually using college students in academic groups to stand in for work groups, such as Hecht et al., 2002, or Upton, 2009), research in actual workplaces is preferred as it is easier to transfer back to organizations for their benefit (Coghlan & Brannick, 2005; Cummings & Worley, 2009). This practice introduces threats to external validity (due to interactions between the causal relationships and units, outcomes, and settings; Heppner, Wampold & Kivlighan, 2008), however – especially when a study is conducted in only one organization – is potential limitations to applicability beyond the organizational context that housed the study; to reduce this methodological concern any finding of significance in an organization should be replicated across other organizations to information on the scope and limitations of generalizability.

Nearly all of the known empirical literature relating team inputs to outputs is quantitative. However, operationalizing many of the variables (either antecedent or outcome) in such a way that comparisons make sense across organizational contexts can be difficult. Consider that Campion et al. (1996)'s replication study of Campion et al. (1993) used four organizations while the original study used one; the measurable work group productivity output variable was omitted in the replication study because equating production from organizations in different industries was irredeemably problematic. It may be that assessing the antecedents and outputs with group work should not be so firmly quantitative; mixed methods approaches (such as action research; Creswell 2012) would likely explore causal relationships better, though it would only exacerbate the

problem of adequate sample sizes, as survey research tends to be easier to administer to large numbers of respondents than qualitative methods. The pressure for organizational research to explain work group outcomes in terms of quantitative output is understandable, however, as research that explains outputs in terms of abstract constructs is much less compelling for management than research that makes a more explicit business case (Alagaraja, 2013). An additional methodological problem for organizational studies at the work group level is the difficulty of identifying and evaluating enough teams to have a large enough sample size for many statistical methods (McIntyre & Salas, 1995), which moving towards a more mixed-methods approach would only exacerbate.

Data Analysis

This study is correlational: SVO and group potency data will be used as predictor variables, at the center level, and compared with that center's recent production. While group potency and production are continuous variables, SVO measures are categorical. This study uses proportions of pro-socialness at the recruiting center level to transform this categorical individual variable to a group-level continuous variable.

Degrees of freedom. Even though the survey solicits responses from over 1700 recruiters, the data is being consolidated and aggregated at the Recruiting center level; the Third Recruiting Brigade currently has 275 Recruiting centers. For this reason, n=275 rather than the number of completed surveys.

Software. As discussed above, data collection is conducted using Electronic Feedback Management software (Verint, 2014) and exported to comma-delineated files. Data sets from USAREC's data warehouse are similarly be exported to comma-

delineated files. Initial data cleaning and structuring will be conducted using Microsoft

Excel. All statistical tests and graph generation will be conducted using appropriate

packages in R (such as the *nlme* package and the *multilevel* package; Bliese, 2016).

Several appendices (Appendices B through F) are be included showing the syntax

used for these analyses. Each hypothesis has its own appendix and the syntax is captured

and preserved in these appendices.

Multi-level modeling. One assumption that seems implicit in all work group

studies reviewed in chapter two is that predictor variables are all at the work group level.

Based on the hierarchical nature of organizations generally, and the U.S. Army in

particular, it makes sense to consider whether significant variance between production in

work groups is attributable to larger group membership. An analysis using multi-level

modeling is conducted to analyze how much of the total variance between work groups is

not located at the work group level, possibly indicating the need to include explanatory

variables that exist at higher levels of an organization as well as to explore potential

cross-level interactions (Raudenbush & Bryk, 2002). This is done using intraclass

correlation coefficients (Castro, 2002; Shieh, 2012; Shrout & Fleiss, 1979). Table 1

clarifies the hierarchy of the organization, the variables, and the level of analysis of each

variable.

Table 1

Association between Variables and Organizational Heirarchy

Organizational Hierarchy			Variables		
Echelon	Level of Analysis	*n*	Correlates	Predictors	Outcomes
Recruiting Battalion	3	8	Propensity	n/a	n/a
Recruiting Company	2	54	Unemployment	n/a	n/a
Recruiting Center	1	275	Number of recruiters, population	Job design (self-management, participation, task variety, task significance), interdependence, relative size, training, process (potency, social support, workload sharing, communication within-group)	Productivity, manager judgement

Note. This table clarifies that all variables from the CWGEM as well as SVO are located at level one. The correlates are located at the lowest level they can be reliably estimated.

Assessing psychometric properties. Evidence of reliability is generated from the scores collected. In addition to the just-discussed r_{wg}*, Cronbach's α is calculated and reported for each subscale. For the Triple-Dominance SVO measure, the item responses are categorical so no Cronbach's α can be calculated because for α, an item-level variance score is required, which is impossible for categorical data. However, per Van Lange, et al.'s (1997) instructions, the six of the nine items must be consistent to award an SVO label to a respondent. This study uses the standard published by Van Lange, et al., who estimate that at least 80% of respondents are consistent.

Summary

In the context of the proposed study's organizational context, four hypotheses are evaluated. First, it is hypothesized that organizationally-relevant, context-dependent predictors (i.e. the "known" recruiting-specific predictors of performance) will explain some differences between recruiting centers.

- Hypothesis One: A significant amount of variance in each dependent variable will be accounted for by recruiting-specific predictors.

After controlling for these organizationally-relevant predictors, proposition three is investigated. Because the complete CGWEM (in either the 1993 or 1996 form) has never been replicated, this proposition will rather be investigated through a more methodologically rigorous analysis of the factors that are retained. Recall that Campion et al. (1993; 1996) only used bivariate, pairwise comparisons for statistical significance testing; the revised model in this study assesses all retained elements together after accounting for organizational-specific predictors.

- Hypothesis Two: the revised work group effectiveness model will significantly account for variance in recruiting center output over and above recruiting-specific predictors.

Finally, propositions one and two are considered.

- Hypothesis Three: SVO and group potency together will significantly account for at least as much variance in recruiting center output over and above recruiting-specific predictors as does the CGWEM's process theme (including group potency).

- Hypothesis Four: SVO will significantly account for some variance in recruiting center output over and above recruiting-specific predictors and all included elements of the CWGEM retained in the revised model.

This chapter explained key concepts, organizational context, population and sampling, measures, research design and procedures. Given these measures and statistical methods, this study is likely to facilitate effective data gathering and management in order to permit analysis that can answer the research questions and propositions.

CHAPTER IV

RESULTS

This study tested four hypotheses using multilevel modeling (MLM) on a data set consisting of survey responses from Army recruiters matched with performance and population metrics from the recruiting centers these Army recruiters are assigned to. Testing each hypothesis for each of the two dependent variables (group productivity and managerial judgement) provided the following insights:

- Hypothesis One: the null hypothesis is rejected for group productivity but not rejected for managerial judgement. While some organizationally-valued and relevant predictors are significantly related to work group productivity, none are related to managerial judgement.

- Hypothesis Two: the null hypothesis is rejected for both group productivity and managerial judgement. The revised work group effectiveness model developed in this paper is related to both managerial judgement and productivity (hypothesis two) although only a few of the predictor variables are individually significant.

- Hypothesis Three: the null hypothesis is rejected for group productivity but not rejected for managerial judgement. Group potency and SVO present as good an explanation of productivity as the process theme from the CWGEM, although in the case of managerial judgement, the original process theme was superior.

80

- Hypothesis Four: the null hypothesis is rejected for group productivity but not rejected for managerial judgement. For group productivity, SVO significantly explains differences in outcomes between groups over and above the contributions of organizationally-relevant predictors and the CWGEM, but does not for the managerial judgement outcome.

Several additional results bear noting at the outset of this chapter. A secondary estimation technique – Markov Chain Monte Carlo (MCMC) simulation – is used to protect against biased estimates based on the choice of analytic tools and to provide evidence that for the MLM parameter estimates. Social Value Orientation assessment of U.S. military personnel is not known to have been previously conducted or reported, so a descriptive comparison between this sample and results from other empirical assessments of SVO in adult populations is included. Aggregating individual SVO assessments as a collective measure (through group proportion of the population-dominant orientation – prosocial) provides a useful explanatory predictor.

This chapter is structured and subdivided using the following sections: survey administration, analytic choices and decisions, hypothesis tests, and summary. The survey administration section reports individual and group descriptives and in particular discusses the SVO results. The analytic choices and decisions section clarifying particular choices and decisions for conducting MLM, lists key R scripts used in this chapter and includes a discussion on centering and scaling. The hypothesis tests section develops a null (predictor-less) model for each dependent variable and then moves through the four hypothesized relationships; each hypothesis subsection includes models

with all hypothesized predictors, a limited model with only significant predictors, and model fit tests, as well as simulation-based secondary estimation.

Survey Administration

The survey was distributed online to subjects' work email addresses via an in-house server running Enterprise Feedback Management (EFM) software (Verint, 2014). The survey was live for 15 days and each potential candidate was contacted a maximum of three times: an initial invitation and survey link on July 6th, 2017, with two follow-up reminders to those who had not yet completed the survey. The follow-ups were sent one and two weeks, respectively, after the initial invitation (on July 13th and 20th); additionally, the second invitation noted that there was only one day left. The survey closed after business hours the evening of July 21st. Of the 1507 Army recruiters in the sample population, 767 completed surveys were returned (individual response rate of 50.9%). The responses were not uniformly distributed across all 273 recruiting centers, so, as such, 94 centers did not have enough responses the meet inclusion criteria (at least two of five or below, three of six or seven, four of eight or nine, and five of 10 or more), resulting in a recruiting center (or work group) response rate of 65.6% (179 of 273). These inclusion criteria are a realization that every group will have missing data (i.e. most groups did not have 100% response rate). Newman and Sin (2009) note that researchers often exclude groups that do not have enough respondents (called *truncation*), although there does not appear to be generalizable best practices on cutoff criteria available; Newman and Sin also note that this practice has the effect of biasing estimates of lower-level ICC downward, meaning that the reported percentages of variance at the

recruiting center level might be higher had groups with unacceptably low response rates not been excluded.

Table 2

Survey Responses (Individual Descriptives)

Measure	n	Items	M (SD)	α	median r_{wg} *
Relative Size	775	1	2.36 (0.87)	n/a	n/a
Group Potency	805	8	3.96 (0.75)	0.91	0.94
Self-Management	805	3	3.01 (0.88)	0.73	0.73
Participation	792	3	3.58 (1.00)	0.91	0.72
Task Variety	792	3	3.73 (0.81)	0.75	0.84
Task Significance	791	3	3.76 (0.87)	0.79	0.79
Interdependence	776	9	3.56 (0.58)	0.74	0.89
Training	774	3	3.64 (0.95)	0.89	0.83
Social Support	774	3	3.89 (0.82)	0.82	0.80
Workload Sharing	774	3	3.34 (1.05)	0.85	0.68
Cooperation	774	3	3.92 (0.86)	0.86	0.84

Individual responses. Generally, the measures adopted from Campion et al. (1993) performed well (Table 2). All subscales had full coverage (responses from 1 through 5 on the 5-point Likert-Type scales). The respondents' responses were internally consistent within subscales, with Cronbach's α scores ranging between .73 (the self-management subscale was the lowest) and .91 (the participation subscale was the highest); these values indicate modest to high levels of within-person consistency.

These responses also generally demonstrated adequate levels of inter-rater reliability. Median r_{wg}* scores (for eligible groups with at least five respondents; Bliese 2016) ranged between .72 (the participation subscale was the lowest) and .84 (the task

variety subscale was the highest); these values provide evidence supporting the aggregation of individual scores to group level and indicate modest to substantial levels of within-group, between-rater reliability of scores. Castro (2002) notes that traditionally interrater reliability scores of .70 have been viewed as sufficient. However, Smith-Crow, Burke, Cohen, and Doveh (2014) determined that the critical values for inter-rater reliability measures are functions of the number of responses per item, number of raters in a group, the correlation between items in a given scale (ρ), and estimated null distribution of responses. The three subscales with the lowest interrater reliability estimates (self-management, participation, and workload sharing) fall a bit short of these critical value estimates: based on Smith-Crow et al.'s work critical values between .75 and .80 should provide reasonable evidence supporting aggregation. Guzzo et al. (1993)'s group potency measure demonstrated similarly desirable results ($\alpha = .91$; median $r_{wg}^{*} = .94$). The relative size measure is a one-item measure (following Campion et al.), so neither internal consistency nor internal reliability ratings are reported. With the exception of the relative size measure, all subscales were modestly but significantly correlated with each other (Table 3).

Group responses. The individual scores from Army recruiters were aggregated to their respective recruiting centers, creating the group-level variables used in this study. These variables, as well as center-level performance metrics (the dependent variables in this study), and company- and battalion-level variables are reported in Table 4. The correlations of the dependent variables and the center-level performance metrics are presented as well (Table 5). There is a high degree of correlation between many of the organizational predictors, but this is expected in organizational research because the

theoretical underpinnings of organizational behaviors and outcomes are often highly related (Heydebrand 1973).

Table 3

Summary of Intercorrelations for Individual Responses

Measure	1	2	3	4	5	6	7	8	9	10	11
1. Relative Size	1.00										
2. Group Potency	0.04	1.00									
3. Self Management	0.07	0.31*	1.00								
4. Participation	0.04	0.60**	0.48**	1.00							
5. Task Variety	0.07	0.64**	0.29**	0.73*	1.00						
6. Task Significance	0.06	0.76**	0.28**	0.66**	0.72*	1.00					
7. Interdependence	-0.02	0.47*	0.23*	0.44**	0.53*	0.49*	1.00				
8. Training	0.04	0.61*	0.25*	0.60**	0.65*	0.64**	0.40*	1.00			
9. Social Support	0.09**	0.67*	0.31*	0.65**	0.64*	0.73**	0.50*	0.69*	1.00		
10. Workload Sharing	0.08****	0.49*	0.24**	0.44**	0.46*	0.49*	0.32*	0.54**	0.58*	1.00	
11. Cooperation	0.06	0.65*	0.26*	0.61*	0.66*	0.65*	0.51*	0.67**	0.85*	0.63*	1.00

Note. *** $p < .05$; ** $p < .01$; * $p < .001$

85

Table 4

Descriptive Statistics of Variables in Study

Variable	*n*	min	max	*M* (*SD*)
Recruiting Center Outcome Variables	179			
Production per Required Recruiter		2.33	12.46	5.95 (1.87)
Mission Accomplishment (% * 100)		36.8	142.9	86.81 (19.67)
Recruiting Center Predictor Variables	179			
Required Recruiters		2.00	14.00	6.70 (2.74)
Population (1000s)		3.60	90.45	25.62 (15.04)
Relative Size		1.00	4.00	2.35 (0.59)
Group Potency		2.57	5.00	3.97 (0.52)
Self-Management		1.33	4.83	3.03 (0.59)
Participation		1.50	5.00	3.60 (0.68)
Task Variety		2.08	5.00	3.76 (0.53)
Task Significance		2.17	5.00	3.81 (0.57)
Interdependence		2.67	4.67	3.58 (0.33)
Training		1.50	5.00	3.67 (0.64)
Social Support		2.00	5.00	3.93 (0.54)
Workload Sharing		1.60	5.00	3.39 (0.74)
Cooperation		1.50	5.00	3.96 (0.61)
Pro-Social SVO Proportion		0.00	1.00	0.79 (0.24)
Recruiting Company Predictor Variable	54			
Unemployment (% * 100)		3.45	7.73	5.26 (1.07)
Recruiting Battalion Predictor Variable	8			
Propensity (% * 100)		5.25	11.07	8.43 (1.82)

Note. For scores not derived from Likert-type scales or unitary counting, parentetical notes indicate units (specifically converting percentages from decimals to ranges from zero to 100, and dividing population by 1000).

Table 5

Summary of Intercorrelations for Recruiting Center Outcomes and Predictors

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Production per Required Recruiter	1.00															
2. Mission Accomplishment	0.77*	1.00														
3. Required Recruiters	-0.17*	0.00	1.00													
4. Population	-0.16***	-0.07	0.74*	1.00												
5. Relative Size	0.11	0.14	-0.15	-0.26*	1.00											
6. Group Potency	0.40*	0.45*	-0.15***	-0.12	0.10	1.00										
7. Self-Management	0.20**	0.23**	-0.16***	-0.20**	0.14	0.48*	1.00									
8. Participation	0.23**	0.22**	-0.27**	-0.23**	0.13	0.65*	0.70*	1.00								
9. Task Variety	0.25*	0.25*	-0.29*	-0.22**	0.12	0.71*	0.45*	0.81*	1.00							
10. Task Significance	0.35*	0.36*	-0.24**	-0.20**	0.14	0.84*	0.48*	0.69*	0.78*	1.00						
11. Interdependence	0.18***	0.14	-0.23**	-0.17***	0.06	0.51*	0.34*	0.47*	0.59*	0.50*	1.00					
12. Training	0.25*	0.22**	-0.25*	-0.19***	0.09	0.67*	0.42*	0.67*	0.77*	0.69*	0.52*	1.00				
13. Social Support	0.22**	0.2**	-0.26*	-0.26*	0.19***	0.69*	0.49*	0.72*	0.75*	0.75*	0.49*	0.79*	1.00			
14. Workload Sharing	0.16***	0.11	-0.34*	-0.28*	0.16***	0.57*	0.35*	0.59*	0.62*	0.57*	0.40*	0.62*	0.70*	1.00		
15. Cooperation	0.14	0.14	-0.23**	-0.21**	0.13	0.66*	0.42*	0.68*	0.75*	0.67*	0.53*	0.78*	0.91*	0.74*	1.00	
16. Proportion of Pro-Socials (SVO)	-0.02	0.06	0.03	0.06	-0.13	0.26*	0.11	0.21**	0.14	0.16***	-0.01	0.15***	0.08	0.15***	0.12	1.00

Note. *** $p < .05$; ** $p < .01$; * $p < .001$

Social value orientation. The SVO triple dominance measure performed better than is routinely reported in the literature. This instrument asks respondents nine questions each consisting of choices differing in outcomes between what one scores and one's partner/opponent scores; each question has three choices, one where the total sum outcomes is maximized (prosocial), one where one's own outcome is maximized (individualist), and one where the differential between outcomes benefits oneself the most (competitive). While an alpha score is not reported for this scale (due to the incompatibility between alpha or other common coefficients and scales with multiple categorical responses), there is a minimum level of internal consistency built into the scale: assigning someone an SVO trait with the triple dominance measure requires that at least six of the nine items indicate the same preference in outcome.

Over 93% of the 768 SVO responses were successfully classified: 606 (78.9%) are prosocial, 82 (10.7%) are individualist, and 28 (3.6%) are competitors, while 52 (6.8%) were unclassified (lacking at least 6 responses indicating any one preference). This is a much higher classification rate than is normal for this instrument: Murphy and Ackermann (2014) review multiple studies that report a failure-to-classify rate of between 25% and 30% - roughly four times as high as the 6.8% in this study. Of those that are successfully classified, 84.6% were prosocial (606 of 716), 11.5% were individualists, and 3.9% were competitors. These results are surprising, as Van Lange et al. (1997) report that, for the age range of subjects in the present sample, prosocials generally account for between 55 and 65% of a given sample, individualists between 25 and 30%, and competitors between 10 and 15%.

The individual SVO responses were converted to work-group scores. To aggregate these individual categorical classifications to the work group level, each work group was assessed in terms of the proportion of its members having the population-dominant SVO trait of prosocial. This distribution is indicated in table 8 and its correlations with other variables in Table 9.

Analytic Choices and Decisions

The planned analysis of this data set was to use three-level MLM incorporating the recommendations of McNeish and Wentzel (2017). Specifically, the recommendation is to use restricted maximum likelihood with a Kenward-Roger degrees-of-freedom approximation and a fixed-effect-only approach to produce estimates of level 1 and level 2 parameters that are fairly unbiased. Unfortunately, because of the reduction in available degrees of freedom at level 1 (the work group level) due to a third of the recruiting centers not having enough responses, the multilevel modeling only considers variance at the recruiting center and the recruiting company level (i.e. a two-level model). The recommendations are the same for conducting MLM with small sample sizes using a two-level model: McNeish and Stapleton (2014) conclude that "with continuous outcomes especially, REML estimation is universally preferable to FML for unbiased variance component estimation," and also recommend including the Kenward-Roger approximation to mitigate potential alpha inflation. Finally, McNeish and Stapleton argue that – for cases with small sample sizes such as the present study – a non-MLM approach such as MCMC is preferable, providing unbiased estimates with smaller degrees of freedom. This study follows these recommendations, using the restricted maximum likelihood and the Kenward-Roger approximation when comparing models.

Additionally, MCMC simulations are reported for each accepted model, and the simulated parameters are reported alongside MLM estimates, allowing the reader to assess the degree in which estimates are biased.

All analysis was conducted using R Studio (R Core Team, 2015). The following packages were used:

- Multilevel models constructed using the *lmer* command in the *lme4* package (Bates, Mächler, Bolker, & Walker, 2014).

- Comparisons between nested MLM used the *KRmodcomp* command in the *pbkrtest* package (Halekoh & Højsgaard, 2014; Kenward & Roger, 1997).

- Markov Chain Monte Carlo simulations conducted using the *MCMCglmm* command in the *MCMCglmm* package (Hadfield, 2010).

Each hypothesis check consists of model development and specification for each of the two work group outputs. Each hypothesis will be investigated for both dependent variables. In both cases, an unconditional or null model is generated to quantify and locate the variance in dependent variables in the absence of predictors. Then, a grand model with all theorized predictors is tested against the null model, and then a restricted or limited model with only somewhat significant ($\alpha < .10$) predictors are tested against each. In cases where there is not a significant difference, the more parsimonious model (i.e. the one with the fewest estimated parameters) is retained. The testing between model comparisons uses an F test, where the test statistic (F_{TS}) is -2*difference between Bayesian log-likelihood values of the models the numerator degrees of freedom (df_n) is the difference of number of estimated parameters between the two models, the denominator degrees of freedom (df_d) is a conservative estimate of average degrees of

freedom in the models (using the Kenward-Roger approximation for degrees of freedom); this test is conducted using the *KRmodcomp* function of the *pbkrtest* package – the commands and outputs are in the appropriate appendices.

Predictor variables are centered to provide meaningful zero values (Raudenbush & Bryk, 2002): recruiters centered to 7, population centered to 25k, unemployment to 5%, propensity to enlist to 10%, and all survey response scales centered to the scale grand mean. Dependent variables are left uncentered, but the production dependent variable is scaled to a monthly rate for ease of interpretation (so divided by 9 as the production data is Q1-Q3, FY17).

Hypothesis Tests

Null model. To provide a baseline to begin hypothesis testing, predictorless (i.e. null or unconditional) models are generated. These unconditional models (Tables 6 and 7) indicate the estimated variance components for both levels as well as the fixed effect for each of the two outcome variables (productivity and manager judgement). The interclass correlation (Raudenbush & Bryk, 2002) indicated the percentage of variance that is attributable to the recruiting company: for productivity (per recruiter per month), it is .01142 / (.01142+.03269) = .2589 or 25.89%; for managerial judgement (mission accomplishment), it is 29.32 / (29.32 + 358.71) = .07556 or 7.56% (See Appendix 2 for R script and output). In each case, the company that a recruiting center is assigned to matters, although the variance attributable to recruiting companies is much larger in the case of per-recruiter efficiency. One way to interpret this outcome within the organizational context would be: the process used to establish recruiting goals accounts for many of the company-level differences that impact recruiter efficiency.

Table 6

Unconditional Model (Productivity)

Fixed Effect	Estimate	se
Production Per Recruiter, γ_{00}	0.660	0.0202
Random Effect	n	Variance
Recruiting Company effect, u_{0j}	54	0.0114
Recruiting Center effect, r_{ij}	179	0.0327

Table 7

Unconditional Model (Managerial Judgement)

Fixed Effect	Estimate	se
Mission Accomplishment, γ_{00}	86.889	1.614
Random Effect	n	Variance
Recruiting Company effect, u_{0j}	54	29.32
Recruiting Center effect, r_{ij}	179	358.71

Hypothesis one. The first hypothesis is designed to check whether the correlates to group productivity that are organizationally-relevant (or contextual) are related to observed outputs (productivity and managerial judgement) – that is to ask – how well do the factors that the organization states are relevant explain the observed outcomes? The null hypothesis is that there is no relationship between these predictor variables and either outcome. Based on the discussion below, the null hypothesis is rejected for the productivity outcome (that is to say, a model using organizationally-valued predictors is significantly related to outcomes) but is not rejected for the managerial judgement outcome (no model using organizationally-valued predictors is significantly related to

mission accomplishment). The "best" model for recruiting center productivity after considering these organizationally-valued predictors for productivity is the limited model (Table 9) and the null model for managerial judgement (Table 7). Of the variance identified in the null model, the limited model for productivity explained 13.24% (unexplained variance reduced from .0114 to .0099) of the variance at the recruiting company level and only 1.03% (from .0327 to .0324) of the variance at the recruiting center level. All scripting and outputs related to testing hypothesis one are located in Appendix 3.

Productivity. Both the grand model (Table 8) and limited model (Table 9) were significantly better models than the null model at predicting productivity (F_{TS}=2.29, df_n=4, df_d=96.6, p=.07 and F_{TS}=3.71, df_n=2, df_d=93.4, p=.03, respectively). The limited model – which retains the number of required recruiters in each recruiting center and the adult unemployment rate of the recruiting center as significant predictors – is simpler yet not statistically different (F_{TS}=.87, df_n=2, df_d=93.0, p=.42) than the grand model, so is selected.

Table 8

Organizational Predictors Grand Model (Productivity)

Fixed Effect	Estimate	se	df	t (p)
Production Per Recruiter, γ_{00}	0.675	0.0275	46.7	24.59 (<.01)
Required Recruiters, γ_{01}	-0.010	0.0084	173.2	-1.21 (.23)
Population, γ_{02}	0.000	0.0016	170.2	-0.03 (.98)
Propensity, γ_{03}	0.014	0.0108	52.4	1.32 (.19)
Unemployment, γ_{04}	0.025	0.0192	51.1	1.32 (.19)
Random Effect	n	Variance		
Recruiting Company effect, u_{0j}	54	0.0099		
Recruiting Center effect, r_{ij}	179	0.0324		

Table 9

Organizational Predictors Limited Model (Productivity)

Fixed Effect	Estimate	se	df	t (p)
Production Per Recruiter, γ_{00}	0.65	0.0199	46.5	32.64 (<.01)
Required Recruiters, γ_{01}	-0.01	0.0056	171.2	-1.79 (.07)
Unemployment, γ_{02}	0.032	0.0183	49.5	1.74 (.09)
Random Effect	n	Variance		
Recruiting Company effect, u_{0j}	54	0.0099		
Recruiting Center effect, r_{ij}	179	0.0324		

The two retained correlates share a great amount of variance with the two correlates that were removed from the model: required recruiters are assigned to recruiting centers based on the local population, so it makes sense that there is a high degree of covariance between the two. In a similar (although lesser) manner, the propensity to enlist is correlated with the local unemployment rate. A "better" – or more relevant to potential military recruits – measure of unemployment would likely be a projected youth unemployment rate (which is not available from the BLS), but even so, the U3 measure accounts for most of the variance that propensity does, and is available as a more granular level (company vs. battalion).

Managerial judgement. The method that USAREC uses to assign monthly missions to its recruiting centers adequately account for these four suspected correlates to work group outcomes. No combination of these variables is statistically better than the null model; the grand model (Table 10) with all four predictors was not statistically

significant (F_{TS}=1.06, df_n=4, df_d=86.0, p=.38). As a result, the null model will be retained and the null hypothesis is not rejected for managerial judgement.

Table 10

Organizational Predictors Grand Model (Managerial Judgement)

Fixed Effect	Estimate	se	df	t (p)
Production Per Recruiter, γ_{00}	85.177	2.2513	41.6	37.84 (<.01)
Required Recruiters, γ_{01}	0.898	0.8151	159.6	1.10 (.27)
Population, γ_{02}	-0.155	0.1527	122.4	-1.01 (.31)
Propensity, γ_{03}	-0.874	0.8977	50.6	-0.94 (.34)
Unemployment, γ_{04}	2.516	1.5892	45.2	1.58 (.12)
Random Effect	n	Variance		
Recruiting Company effect, u_{0j}	54	23.1		
Recruiting Center effect, r_{ij}	179	363.2		

Secondary estimations. The parameters in the accepted model are assessed for plausibility using a secondary estimation technique. Markov Chain Monte Carlo (MCMC)-based parameter estimates for the model are calculated and compared to the original multilevel model-based estimates. This simulation had a long burn-in (2000 iterations) and a long runtime (28,000 iterations beyond burn-in). Table 11 shows that the estimates MLM estimates presented above (for parameter estimates, calculated probability values, and level-based variance) are reasonable and that biased estimates due to statistical method is unlikely.

Table 11

Comparison of MCMC and MLM Estimates (Productivity)

	MCMC			MLM	
Fixed Effect	Estimate	*95% CI (lower, upper)*	*p*	Estimate	*p*
Production Per Recruiter, γ_{00}	0.65	(0.61, 0.69)	<.01	0.65	<.01
Required Recruiters, γ_{01}	-0.01	(-0.02, 0.00)	.07	-0.01	.07
Unemployment, γ_{02}	0.03	(-0.00, 0.07)	.09	0.03	.09
	MCMC			MLM	
Random Effect	Variance	*95% CI (lower, upper)*		Estimate	
Recruiting Company effect, u_{0j}	0.0090	(0.0000, 0.0177)		0.0099	
Recruiting Center effect, r_{ij}	0.0340	(0.0252, 0.0437)		0.0324	

Note. The MLM values are repeated from Table 9 and presented here for convenience.

Hypothesis Two. The second hypothesis is that the predictors from the revised work group effectiveness model will significantly account for variance in work group outputs over and above context-specific predictors (which were tested and explored as hypothesis one). The null hypothesis two (that there is no predictive relationship between constructs in the CWGEM and productivity or managerial judgement) is rejected for productivity; both a model with all 11 predictors as well as a more limited model of two statistically significant predictors – group potency and cooperation & communication within-group – are significantly better than the model from hypothesis one. This limited model for productivity reduces the work group-level variance to .0295 from the best hypothesis one model (.0324), a reduction of 8.76% of the residual variance and a cumulative reduction of 9.71% of the initial variance as calculated in the null model.

Similarly, the grand and limited models for managerial judgement were both statistically significant when compared to the null model (which was retained after hypothesis one). They were not statistically different from each other, so the limited

model which retained three predictors (group potency, cooperation & communication within group, and relative size) is selected and retained. This limited model for managerial judgement reduces the work group-level variance by 21.5% (from 358.71 to 281.65).

Productivity. Including all 11 predictor variables from the CWGEM that were assessed through surveying recruiters resulted in significantly better model than the best model retained in hypothesis one (F_{TS}=2.97, df_n=11, df_d=159.4, p<.01); Table 12 reports all estimated parameters. Many of the variables do not appear to be significant; however, so they are removed one at a time (highest p value each time) until all remaining variables have a p of less than .10. This results in a limited model (retaining two of the subscales from the revised CWGEM: group potency and cooperation & communication within group) that is significantly better than the best model retained in hypothesis one (F_{TS}=15.26, df_n=2, df_d=171.5, p<.01) and is simpler yet not significantly different than the hypothesis two grand model (F_{TS}=.30, df_n=9, df_d=159.3, p=.97), so is selected as the best current model as hypothesis two's null model is rejected. Table 13 reports this best limited model.

Table 12

Revised CWGEM Grand Model (Productivity)

Fixed Effect	Estimate	se	df	t (p)
Production Per Recruiter, γ_{00}	0.65	0.0171	38.5	38.22 (<.01)
Required Recruiters, γ_{01}	-0.01	0.0059	165.0	-1.59 (.11)
Unemployment, γ_{02}	0.02	0.0163	41.9	1.06 (.30)
Group Potency, γ_{03}	0.17	0.0530	154.4	3.29 (<.01)
Self Management, γ_{04}	0.00	0.0358	162.5	-0.09 (.93)
Participation, γ_{06}	0.00	0.0478	164.5	0.08 (.94)
Task Variety, γ_{07}	-0.01	0.0650	164.0	-0.08 (.94)
Task Significance, γ_{08}	0.01	0.0559	151.3	0.10 (.92)
Interdependence, γ_{09}	-0.01	0.0541	153.6	-0.15 (.88)
Relative Size, γ_{10}	0.02	0.0250	162.3	0.64 (.52)
Training, γ_{11}	0.04	0.0404	152.9	0.96 (.34)
Workload Sharing, γ_{12}	-0.01	0.0303	163.1	-0.48 (.63)
Social Support, γ_{13}	0.05	0.0767	161.6	0.66 (.51)
Cooperation, γ_{14}	-0.13	0.0667	165.0	-1.94 (.05)

Random Effect	n	Variance		
Recruiting Company effect, u_{0j}	54	0.0050		
Recruiting Center effect, r_{ij}	179	0.0312		

Table 13

Revised CWGEM Limited Model (Productivity)

Fixed Effect	Estimate	*se*	*df*	*t (p)*
Production Per Recruiter, γ_{00}	0.65	0.0172	39.8	38.01 (<.01)
Required Recruiters, γ_{01}	-0.01	0.0053	172.2	-1.88 (.06)
Unemployment, γ_{02}	0.02	0.0160	42.9	1.17 (.02)
Group Potency, γ_{03}	0.19	0.0354	170.9	5.46 (<.01)
Cooperation, γ_{04}	-0.08	0.0305	168.4	-2.62 (.01)
Random Effect	n	Variance		
Recruiting Company effect, u_{0j}	54	0.0057		
Recruiting Center effect, r_{ij}	179	0.0295		

Managerial judgement. Including all 11 predictor variables from the CWGEM that were assessed through surveying recruiters resulted in significantly better model than the best model retained in hypothesis one (F_{TS}=5.33, df_n=11, df_d=164.0, p<.01); Table 14 reports all estimated parameters. Many of the variables do not appear to be significant; however, so they are removed one at a time (highest p value each time) until all remaining variables have a p of less than .10. This results in a limited model (retaining three of the subscales from the revised CWGEM: group potency, cooperation & communication within group, and relative size) that is significantly better than the best model retained in hypothesis one (F_{TS}=19.22, df_n=3, df_d=173.9, p<.01) and is simpler yet not significantly different than the hypothesis two grand model (F_{TS}=.30, df_n=7, df_d=165.4, p=.95), so is selected as the best current model as hypothesis two's null model is rejected. Table 15 reports this best limited model.

Table 14

Revised CWGEM Grand Model (Managerial Judgement)

Fixed Effect	Estimate	se	df	t (p)
Mission Accomplishment, γ_{00}	86.84	1.3509	41.5	64.28 (<.01)
Group Potency, γ_{02}	22.36	4.8881	164.5	4.57 (<.01)
Self Management, γ_{03}	2.68	3.3224	166.9	0.81 (.42)
Participation, γ_{04}	-2.90	4.3887	164.3	-0.66 (.51)
Task Variety, γ_{05}	2.53	5.9877	166.3	0.42 (.67)
Task Significance, γ_{06}	1.28	5.2320	163.3	0.24 (.81)
Interdependence, γ_{07}	-4.01	5.0082	160.7	-0.80 (.42)
Relative Size, γ_{08}	4.13	2.2977	166.2	1.80 (.07)
Training, γ_{09}	1.43	3.7805	165.8	0.38 (.71)
Workload Sharing, γ_{10}	-2.90	2.7094	167.0	-1.07 (.29)
Social Support, γ_{11}	-0.54	7.1193	166.5	-0.08 (.94)
Cooperation, γ_{12}	-6.97	6.0714	166.0	-1.15 (.25)

Random Effect	n	Variance
Recruiting Company effect, u_{0j}	54	8.74
Recruiting Center effect, r_{ij}	179	294.05

Table 15

Revised CWGEM Limited Model (Managerial Judgement)

Fixed Effect	Estimate	se	df	t (p)
Mission Accomplishment, γ_{00}	86.85	1.3510	44.2	64.30 (<.01)
Group Potency, γ_{01}	23.24	3.2560	174.2	7.14 (<.01)
Relative Size, γ_{02}	4.12	2.2060	173.1	1.87 (.06)
Cooperation, γ_{03}	-9.19	2.8000	175.0	-3.28 (<.01)

Random Effect	n	Variance
Recruiting Company effect, u_{0j}	54	12.11
Recruiting Center effect, r_{ij}	179	281.65

Secondary estimations. The parameters in the accepted model are assessed for plausibility using a secondary estimation technique. Markov Chain Monte Carlo (MCMC)-based parameter estimates for the model is calculated and compared to the original multilevel model-based estimates. This simulation had a long burn-in (2000 iterations) and a long runtime (28,000 iterations beyond burn-in). Tables 16 and 17 shows that the estimates MLM estimates presented above (for parameter estimates, calculated probability values, and level-based variance) are reasonable and that biased estimates due to statistical method is unlikely.

Table 16

Comparison of MCMC and MLM Estimates (Productivity)

	MCMC			MLM	
Fixed Effect	Estimate	*95% CI (lower, upper)*	*p*	Estimate	*p*
Production Per Recruiter, γ_{00}	0.65	(0.62, 0.68)	<.01	0.65	<.01
Required Recruiters, γ_{01}	-0.01	(-0.02, 0.00)	.07	-0.01	.06
Unemployment, γ_{02}	0.02	(-0.01, 0.05)	.26	0.02	.25
Group Potency, γ_{03}	0.20	(0.13, 0.27)	<.01	0.19	<.01
Cooperation, γ_{04}	-0.08	(-0.14, -0.02)	.01	-0.08	.01

	MCMC		MLM
Random Effect	Variance	*95% CI (lower, upper)*	Estimate
Recruiting Company effect, u_{0j}	0.0029	(.0000, .0101)	0.0057
Recruiting Center effect, r_{ij}	0.0326	(.0243, .0415)	0.0295

Note. The MLM values are repeated from Table 13 and presented here for convenience.

Table 17

Comparison of MCMC and MLM Estimates (Managerial Judgement)

Fixed Effect	MCMC			MLM	
	Estimate	*95% CI (lower, upper)*	*p*	Estimate	*p*
Mission Accomplishment, γ_{00}	86.81	(84.29, 89.36)	<.01	86.85	<.01
Relative Size, γ_{01}	4.07	(-0.27, 8.36)	.07	4.12	.06
Group Potency, γ_{03}	23.21	(16.92, 29.71)	<.01	23.24	<.01
Cooperation, γ_{04}	-9.03	(-14.49, -3.28)	<.01	-9.19	<.01
Random Effect	MCMC			MLM	
	Variance	*95% CI (lower, upper)*		Estimate	
Recruiting Company effect, u_{0j}	<0.01	(<0.01, <0.01)		12.11	
Recruiting Center effect, r_{ij}	297.20	(234.80, 359.00)		281.65	

Note. The MLM values are repeated from Table 15 and presented here for convenience.

Hypothesis Three. As hypothesis two built on the results from hypothesis one, so does this hypothesis, which considers the best way to specify the process theme that Campion et al. (1993; 1996) and other scholars have found compelling and predictive, in that group potency and SVO will be at least as good as predictors of group outcomes than group potency, social support, workload sharing, and cooperation & communication within-group. As such, two new models will are specified and compared for each outcome (productivity and managerial judgement). The null hypothesis (that the theorized replacement model including SVO's predictive ability is inferior to the existing CWGEM's process subtheme) is rejected for productivity, but is not rejected for managerial judgement.

Because the models are not nested (in that one is not a simpler version of the other), the conservative KR test cannot be used, so rather the Akaike Information Criterion, or AIC, (Akaike, 1973) will be used for model selection (Pawitan, 2001).

Model selection with AIC recommends models with smaller values (Wagenmakers & Farrell, 2004), with the greater discrepancy interpreted as more evidence to prefer one to another.

Productivity. For productivity, both models are significantly better than the best model from hypothesis one, indicating that both models better explain productivity than the limited context-specific model (F_{TS}=13.35, df_n=2, df_d=171.0, p<.01 and F_{TS}=8.05, df_n=4, df_d=169.9, p<.01 for SVO and the process theme, respectively). The AIC scores are very similar for the SVO model (-83.63) and the CWGEM productivity theme (-84.96), so the simpler model (group potency and SVO) is retained. The null hypothesis is rejected.

Managerial judgement. In terms of managerial judgement, both models are significantly better than the best model from hypothesis one (which was the null model), indicating that both models better explain productivity than the limited context-specific model (F_{TS}=21.55, df_n=2, df_d=174.9, p<.01 and F_{TS}=13.7, df_n=4, df_d=172.2, p<.01 for SVO and the process theme, respectively). Assessing the difference in AIC values, however, the null hypothesis is not rejected because the CWGEM's productivity theme was moderately lower (1537.09) than group potency and SVO (1543.09).

Hypothesis Four. Adding SVO to the best model from hypothesis two is marginally significant for productivity (F_{TS}=3.64, df_n=1, df_d=166.2, p=.06) but not for managerial judgement (F_{TS}=.49, df_n=1, df_d=172.3, p=.48). As a result, the null hypothesis is rejected for productivity (i.e. SVO appears to account for significant variance in productivity over and above the earlier variables included in the model) but not rejected

for managerial judgement. The estimated parameters for the alternate model for productivity – as well as secondary estimations with MCMC – are reported in Table 18.

Table 18

Comparison of MCMC and MLM Estimates (Productivity)

| Fixed Effect | MCMC | | | MLM | |
	Estimate	*95% CI (lower, upper)*	*p*	Estimate	*p*
Production Per Recruiter, γ_{00}	0.75	(0.65, 0.85)	<.01	0.74	<.01
Required Recruiters, γ_{01}	-0.01	(-0.02, 0.00)	.08	-0.01	.07
Unemployment, γ_{02}	0.01	(-0.01, 0.04)	.33	0.02	.29
Group Potency, γ_{03}	0.23	(0.15, 0.30)	<.01	0.21	<.01
Cooperation, γ_{04}	-0.08	(-0.14, -0.02)	.01	-0.08	<.01
Prosocial SVO Proportion, γ_{05}	-0.12	(-0.24, 0.00)	.05	-0.11	.06

| Random Effect | MCMC | | MLM |
	Variance	*95% CI (lower, upper)*	Estimate
Recruiting Company effect, u_{0j}	0.0005	(0.0000, 0.0043)	0.0053
Recruiting Center effect, r_{ij}	0.0341	(0.0264, 0.0420)	0.0293

Summary

For the work group productivity outcome, significant relationships were found for all four hypotheses; resulting in the rejection of the null models. For the managerial judgement outcome, most of the hypothesized relationships were insignificant, so – with the exception of hypothesis two – the null hypotheses were not rejected.

The difference in outcomes between the two dependent variables will be discussed in some detail in the following chapter, though – briefly – the likely explanation for the difference between the two is the formalized process of goal-setting used by the studied organization, which takes past performance into account, and predictors that may change very slowly over time seem to be plausibly accounted for in

the mission model. Simulation-based secondary parameter estimation resulted in values very similar to the restricted likelihood MLM estimates.

CHAPTER V

DISCUSSION

This study was designed to develop and test a revised model that better explains

differences in outputs based on differences in predictive factors work groups and teams.

This was primarily accomplished by reviewing the work by Campion et al. (1993; 1996),

refining the model presented by Campion et al. in light of subsequent research, and

including two additions: a block of organizationally-relevant predictors (i.e. appropriate

to and valued by the studied organization) before testing the CWGEM, and including

SVO as an additional predictor. In doing so, four hypotheses were assessed across two

work group outcomes (productivity and managerial judgement):

- Hypothesis One: that organizationally-relevant (i.e. business- or industry-specific) predictors explain some differences in work group outcomes

- Hypothesis Two: that the revised CWGEM explains some differences in work group outcomes over and above organizationally-relevant predictors

- Hypothesis Three: that SVO and group potency together explain some differences in work group outcomes at least as well as Campion et al.'s process theme (group potency, social support, workload sharing, and cooperation and communication within-group)

- Hypothesis Four: that SVO explains some differences in work group outcomes over and above the CWGEM and organizationally-relevant predictors

The null hypotheses for all four hypothesized relationships were rejected for the productivity outcome, while only hypothesis two's null was rejected for the managerial judgement outcome. These results, summarized in Table 23, provide evidence that context-relevant organizational predictors are related to productivity (but not managerial judgement), that elements of the CWGEM are related to both work group outcomes (productivity and managerial judgement) over and above the context-relevant organizational predictors, that SVO is a significant predictor of productivity (but not managerial judgement) over and above context-relevant and CWGEM predictors, and that group potency and SVO together explain productivity (but not managerial judgement) at least as well as Campion's process theme (group potency, work load sharing, cooperation & communication within-group, and social support).

Table 23

Summary of Hypothesis Test Results

Work Group Outcome	H_1	H_2	H_3	H_4
Productivity	null rejected	null rejected	null rejected	null rejected
Managerial Judgement	null retained	null rejected	null retained	null retained

Consideration of these findings is conducted primarily through a return to the two original research questions posed at the start of this study: Does SVO provide a useful addition to explaining work group outcomes? How can Campion et al.'s (1993; 1996) model be improved? In terms of structure, this discussion is divided into the following sections: research question one, research question two, organizational and methodological implications, generalizability and limitations, and closing commentary.

Research Question One: Is SVO Predictive of Work Group Outcomes?

Social value orientation – quantified at the group level as the percentage of a work-group that assesses as a prosocial on the Triple Dominance Measure – and group potency explain differences in productivity about as good as the process theme in the CWGEM (group potency, work load sharing, cooperation & communication within group, and social support). It appears that this time-stable personality trait is as good a predictor as three time-variant measures of internal group processes. While this means that SVO might be a good substitute for several of Campion et al.'s (1993; 1996) subscales, SVO also appears to explain additional variability that the CWGEM does not.

In this study, social value orientation is found to be significantly related – over and above contextually-relevant predictors and the CGWEM – to productivity, but not to managerial judgement. This demonstrates that differences in SVO proportions between groups might be a useful factor in explaining differences in performance between groups. In this study, the lack of difference in managerial judgement is quite likely for the same reason that most other hypotheses for the managerial judgement outcome were rejected: SVO is a time-stable trait, and even though members of a group change slowly over time, much of the contributions of SVO are accounted for by USAREC's methodical setting of the recruiting goals that they base their managerial judgement on. This bears further discussion.

The way that USAREC in particular – and organizations in general – choose to evaluate and judge their work groups and teams has several significant implications for researchers and practitioners. By leaning so heavily on past performance to set benchmarks that will serve as the basis of managerial judgements, many predictive

factors might be dismissed as irrelevant because slowly-changing predictors will not appear to be significantly related to valued outcomes. The danger of all this is that organizations may miss opportunities to improve systems, and potentially not implement HRD changes designed to increase performance (high-performance work principles, policies, practices, and products; Posthuma, Campion, Masimova, & Campion, 2013). Decisions in group composition and assignment of people to work groups, in terms of functional diversity or personality preferences, will not be deliberately made because they are not obviously relevant to meeting or failing to meet organizational objectives. A counter argument can be made here, however, as the system that the organizations in Campion et al. (1993; 1996) used are seemingly arbitrary and likely not systematic, with each manager making individual assessments of their subordinate teams without any cohesive assessment plan or scheme, which limits the larger organizations' abilities to evaluate and assess teams between different managers. On balance, a deliberate and analytical plan to evaluate and assess work groups and teams might be best, as long as organizations are aware of any potential blind spots and assess – as the present study has – effectiveness and managerial judgement independently.

Recruiting centers that had a higher proportion of pro-social individuals successfully enlisted individuals into the U.S. Army at a lower per-recruiter rate than did work groups that had a lower proportion of pro-social individuals. Based on the best model developed in hypothesis four (Chapter Four, Table 23) groups that had 100% pro-social individuals are expected to recruit about .11 fewer individuals per recruiter per month than groups with no pro-social individuals (which is a fairly large difference in performance, over one-seventh of average monthly production).

Does this mean that individualists and competitors are better workers, or better recruiters, than pro-social individuals? Not at all – recall the discussion in Chapter 2, specifically Stouten et al. (2005), where SVO differences were demonstrated in times of success and failure; or from Upton (2009) who showed different SVOs being the most productive under different incentive schemes; and finally from Bogaert et al. (2012), who argued that SVO interacted with and moderated the effects of well-regarded yet complicated organizational concepts such as climate and commitment. The relevance of the finding in this study is that while SVO is important, its specific effect on outcome is dependent on the organizational context: the sum total effect of incentives, environment, culture, internal business rules, leadership, etc. Context is critically important here, as it is in any consideration of high-performance work practices (Combs, Liu, Hall, & Ketchen, 2006). Any recommendation for change or improvement needs to be considered and assessed in light of other organizational concerns.

For USAREC, the current state is one where pro-social individuals do not appear to be as successful as do pro-self individuals (individualists and competitors). This is a potential area of misalignment, as the stated and articulated values of the U.S. Army (the "seven Army values") include *selfless service* (Headquarters, Department of the Army, 2015). Additionally, because pro-self individuals in USAREC are fairly uncommon (as nearly 85% of those successfully assessed are pro-social), that this 15% minority seems to have an internal competitive advantage means that 85% - the vast majority of individuals – do not. This misalignment between individual preferences and business rules means that the organization is likely curtailing its ability to be as successful as it could be (Swanson & Holton, 2009). However, specific recommendations to change

appear to be an open question; more research is surely needed before a cost-effective recommendation could be made as to exactly what to change to make pro-socials more productive.

Research Question Two: How can the CWGEM be improved?

It should be noted that the findings of Campion et al. (1993) are supported and partially replicated in this study. The relationships in the Campion et al. study were only bivariate correlations; the intercorrelations reported in Campion et al. Table 1, Columns 1-18 are very similar to the corresponding relationships in this study, reported previously in Table 5, Columns 5-15. Additionally, many similarities in relationships between these predictors and the productivity output are apparent when comparing the productivity column of Table 2 (Campion et al.) to this study's Table 5, Column 1. Because of the high degree of intercorrelation between these predictors, this study considers the relationships between outputs and predictive variables while assessing the contributions of predictors in the context of each other, which Campion et al. did not do. Additionally, this study removed some of the variables that did not seem to be empirically supported by subsequent literature, resulting in a revised CWGEM retaining the most promising variables.

The revised CWGEM developed in the literature review chapter was statistically significant for both outcome variables: it explained differences in productivity and managerial judgement better than the context-relevant predictors in hypothesis one. However, a more parsimonious model with only a few of the predictors from the CWGEM is much simpler, yet not statistically different, than the entire revised model; for this reason the limited models are preferred (reported previously in Tables 13 and 16).

Two predictor variables (group potency and within-group cooperation & communication) are significantly related to both outcomes:

- Group potency is positively related to both outcomes. This relationship is expected from the discussion in the literature review, however, due to the lack of previous studies reporting relative effects of different predictors, that group potency is the most important predictor was not an expected result.

- Cooperation and communication within-group is negatively correlated with both outcomes. Campion et al. (1993; 1996) theorize and report a positive relationship (i.e. the more within-group cooperation and communication, the more desired outcomes); this study reports positive – but nonsignificant – bivariate relationships as well. However, when assessed in the context of other predictive variables, and specifically group potency, the relationship between within-group cooperation and communication and the two outcome variables was negative.

One additional predictor was significant for the managerial judgement outcome but not the productivity outcome. Relative size, which was the felt need for more or less people in the workgroup, is related to managerial judgement in the following way: work groups that say they have groups that are about the right size or a bit too large tend to accomplish a larger percentage of their assigned monthly goals than groups that say they do not have enough people and need the size of their group increased. How can this be, when from hypothesis one we know that groups are actually more productive on a per-person basis when they are smaller, and that actual group size and achievement of recruiting goals are not related? One explanation is that groups that struggle with their

goals might use the size of their group as an excuse, and attribute their lack of success to a lack of human capital.

Organizational and Methodological Implications

Reviewing literature specific to the studied organization – a U.S. Army Recruiting brigade – resulted in identifying several likely predictors that should be controlled for (unemployment, propensity to enlist, population and number of recruiters) prior to considering the relationship of variables of interest to outcomes of interest. Two of these organizationally-relevant variables (unemployment and number of recruiters) were significant predictors of relative productivity between recruiting centers. As unemployment goes up, the per-recruiter monthly production tends to increase, which is expected. As the number of recruiters increases, per-recruiter monthly production decreases – potentially due to inefficiency or decreased cohesion and commitment (per the social loafing literature such as Comer, 1995, Mueller, 2011, or Mulvey & Klein, 1998), or perhaps due to the fact that the number of recruiters is correlated with the size of the population (as presented in Table 9) and that the greater number of people a recruiting center has, then there might be more difficulty making meaningful connections with community influencers and gatekeepers, more lengthy drive times so less hours in the day to meet with people, etc. This is to say, while increased numbers of recruiters is significantly negatively correlated with per-recruiter productivity, the relationship may or not be directly causal. Theorizing why the other two organization-specific predictors (propensity and population) were not significant was because they likely explained much of the same variance as the two variables that were retained: unemployment and propensity to enlist both explain the same decision that young adults have in deciding

whether to join the military or not (the choice as articulated by Kleycamp, 2006), while population and number of recruiters are highly correlated due to the fact that USAREC uses population as a significant factor in setting recruiter numbers at recruiting centers.

None of these context-relevant organizational predictors were significantly related to managerial judgement. From the point of view of the studied organization, this is a positive finding: the business rules and processes used to set recruiting goals (the accomplishment of which are the managerial judgement in this organization) are quite formalized and use factors that should account for the contributions of the four hypothesized predictors; USAREC uses past production and population to set future goals, and the four hypothesized predictors change very slowly over time, so any time-stable predictor should get "baked into" future recruiting goals and then not be statistically significant predictors of managerial judgements. In short, if an organization uses past performance to set future performance goals, then factors which are fairly time-stable that are relevant to productivity will not be relevant to managerial judgement; this explains why hypothesis one was rejected for managerial judgement while accepted for productivity.

The predictor that was so effective at explaining both organizational outcomes in this study (group potency) is not something that is included in the U.S. Army's current list of "characteristics of effective teams" (Headquarters, Department of the Army, 2015, Figure 1-12). Similarly, the stated method of measuring *good teamwork* involves the kind of constructs which are quite similar to predictors in most literature on work group and team effectiveness – "identity, cohesion, and climate" – but not any of the actual outcomes of effective collective effort (such as recruits per month or a subjective

assessment of a group's performance). For these reasons, this study offers a new and extended perspective for military units to consider their small groups and teams.

In the introductory chapter and the literature review chapter, we discussed that nearly all empirical work group and team literature fell into one of three groups: either testing one or a few academic concepts in an empirical study, testing a wide-ranging and comprehensive model in an empirical study (such as Campion el al. 1993 and 1996), or testing some organizationally-relevant predictors. The limitations of each of the three approaches seem obvious: the first is too narrow and does not assess variables in the context of other explanatory variables that were not considered, the second might be overbroad and can be unwieldy to routinely administer to organizations, and the third limits generalizability and does not always include the best scholarly knowledge. The present study tries to thread the needle and take the best of all three: including many – but by no means all – of the most likely scholarly constructs related to work group outcomes and considering them only after testing and including relevant organizationally-relevant predictors. This approach seeks to be at once academically rigorous and yet practical and accessible for practitioners. As such, it might be in the spirit of Merton's idea of a *middle-range* theory (Moore, Johns, & Pinder, 1980), which argued for the need to scope theories and models of organizational behavior somewhere between the specific but context-free and the comprehensive yet overbroad. This approach also follows Heydebrand's (1973) advice to avoid – on the one hand – case studies of limited generalizability, and – on the other – overbroad models which lack clarity or context.

Generalizability and Limitations

The findings related to the productivity outcome should be viewed as more generalizable than the findings related to the managerial judgement outcome. This study operationalized productivity very similar to Campion et al. (1993) and other studies that evaluate tangible, countable production. However, the experiences of Campion et al. (1996) are important to remember, that direct comparison of productivity between different kinds of organizations (or industries) is troublesome, as units of production and coefficients of predictor often do not make sense in other organizations. The directional relationships (i.e. positive or negative correlation, or no relationship) for production are – I would argue – quite likely transferable to many large organizations with many work groups that are similar to each other.

The organization in this study (USAREC) has a formalized process for setting performance goals, and the percentage of goal achievement is the way the organization assesses performance of groups at all echelons. For this reason, managerial judgement is not measured in this study as Campion et al. (1993; 1996) did, which was with a survey instrument as an attitudinal subscale. Caution would be advised to generalize these findings (with respect to the managerial judgement outcome) to other organizations unless (a) the process for managerial judgement is formalized, accepted, and understood universally within the organization, and (within-organization) public, and (b) past performance of groups has significant impact on the future goals. If these criteria are met, then perhaps these managerial findings should be considered to have some predictive capacity in other organizations.

Future research recommended into organizations may take two tracks. First, the need for a refined theory of how SVO preferences within work groups interacts with

(potentially through mediation or moderation) other, more commonly-used constructs is

clear. In hand with this theory should be developments of potential organizational

interventions and recommendations of what to change so that workers and groups of a

particular SVO preference might be more successful, satisfied, and productive. Second, a

realization of the practical limitations of sample size in most real-world organizational

settings might recommend against a presumption of appropriateness of MLM approaches

in organizational research studies specifically, and HRD studies generally. McNeish,

Stapleton, and Silverman (2017) argue this passionately and provide recommendations of

other analytic tools that are often more appropriate and robust in situations of limited

sample size or when the analytic rigor of MLM techniques may not be necessary or

desired. Finally, were this study to be reproduced in a different organization, then the

context-specific predictors in this study should be discarded and replaced with the best

available predictors of performance available to the studied organization.

Closing Commentary

In terms of SVO, I will close with a discussion of an argument for measuring and

including value preferences when investigating organizations. England (1975) crafted a

scheme to measure the value systems of managers: some managers are pragmatic,

primarily focusing on matters of success and failure; others are moralistic, primarily

focusing on matters of right and wrong; some are affective, focusing on concerns of the

pleasant versus the unpleasant. With caveats, England demonstrated a relationship

between managerial orientations and collective outcomes. England's assumptions for his

managerial orientation are different than those that SVO uses, but reaches the same broad

conclusions: different orientations can succeed or fail in different contexts, and matching

organizational factors such as size and culture to managerial traits is important to determine success or failure. SVO looks at preferences in outcome balances between self and others, and in practice various SVO preferences are more or less successful based on their contexts.

Importantly England (1975) and Whitely's (1979) research demonstrates that the values individuals in groups hold influences decision-making and group performance, lending credence to the assertion that SVO should influence the actions, performance, decisions, and outcomes of work groups and teams. While SVO and other value-based research such as England's and Whitely's generally provide credence to a greater argument that value-based variables are important to consider, they rely on different theoretical mechanisms. For instance, Zander (1985) argues that conceptualizations of managerial values explains outcomes because England's value preferences limit choices and actions (by restricting consideration of some types of outcomes), which is distinct from the conceptualization of SVO as a preference between multiple achievable outcomes. I contend that SVO is compatible with this larger body of scholarship into the relationship between organizational performance and the values and preferences of organizational members (such as Agle & Caldwell, 1999), and that generally, scholarship supporting research into the relationship between individual values and organizational outcomes should be read as supporting consideration of SVO's impact into group outcomes in the workplace.

The results of this study provide contributions to our understanding of work groups and teams. Primarily, this is accomplished by revising the CWGEM and extending it (with the inclusion of SVO). Methodologically, where Campion et al. (1993;

1996) used bivariate correlations to consider relationships between predictors and outcomes, this study assesses predictors in the context of each other, and accounts (through MLM) for the organizational hierarchical structure of data sets generated from studying multiple work groups within the same organization. Contributions to HRD include a methodology to contextualize studied constructs with variables and concepts valued by studied organizations; evidence that group potency might be far more important than many other commonly-studied predictors; and rationale to include SVO as a variable to consider in future research.

Practically, this study provides the subject organization affirmation supporting the efficacy of its formal processes that set performance goals and form the basis of evaluating and assessing work group performance. Additionally, the results provide insight into factors that differentiate better- and worse-performing work groups, forming the basis of organizational interventions into some of the underperforming groups. This study provides a practical method of bridging the perceived divide within HRD between scholars and practitioners by assessing scholarly-based constructs only in the context of those valued by studied organizations. Finally, this study found that SVO – categorizing preferences in outcomes in terms of utility and interpersonal conflict – is promising in its relationship to organizational outcomes and its compatibility with existing theories and models.

REFERENCES

Agle, B.R., & Caldwell, C.B. (1999). Understanding research on values in business. *Business & Society, 38*(3), 326-387.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In S. Kotz & N.L. Johnson (Eds.), *Breakthroughs in statistics volume I* (pp. 610-624). New York: Springer-Verlag.

Alagaraja, M. (2013). HRD and HRM perspectives on organizational performance: A review of the literature. *Human Resource Development Review, 12*(2), 117-143. doi://10.1177/1534484312450868

Ancona, D.G. (1990). Outward bound: Strategies for team survival in an organization. *Academy of Management Journal, 33*(2), 334-365. doi://10.2307/256328

Asch, B.J., Kilburn, M.R., & Klerman, J.A. (1999). *Attracting College-Bound Youth into the Military.* Santa Monica, CA: RAND.

Bandura, A. (1998). Personal and collective efficacy in human adaptation and change. In J.G. Adair, D. Belanger & K.L. Dion (Eds.), *Advances in psychological science, Vol. 1: Social, personal, and cultural aspects* (pp. 51-71). Hove, England: Taylor & Francis.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi://10.18637

Barnes, C.M., Hollenbeck, J.R., Jundt, D.K., DeRue, D.S., & Harmon, S.J. (2011).

Mixing individual incentives and group incentives: Best of both worlds or social

dilemma? *Journal of Management, 37*(6), 1611-1635.

doi://10.1177/0149206309360845

Barrick, M.R., Stewart, G.L., Neubert, M.J., & Mount, M.K. (1998). Relating member

ability and personality to work-team processes and team effectiveness. *Journal of

Applied Psychology, 83*(3), 377-391. doi://10.1037//0021-9010.83.3.377

Barry, D. (1991). Managing the bossless team: Lessons in distributed leadership.

Organizational Dynamics, 20(1), 31-47. doi://10.1016/0090-2616(91)90081-j

Bliese, P. (2016). *Multilevel modeling in R (2.6)*. Retrieved from https://cran.r-

project.org/doc/contrib/Bliese_Multilevel.pdf

Bogaert, S., Boone, C., & van Witteloostuijn, A. (2012). Social value orientation and

climate strength as moderators of the impact of work group collective climate on

affective commitment. *Journal of Management Studies, 49*(5), 918-944.

doi://10.1111/j.1467-6486.2011.01029.x

Campion, M.A., Medsker, G.J., & Higgs, A.C. (1993). Relations between work group

characteristics and effectiveness: Implications for designing effective work

groups. *Personnel Psychology, 46*, 823-850. doi://10.1111/j.1744-

6570.1993.tb01571.x

Campion, M.A., Papper, E.M., & Medsker, G.J. (1996). Relations between work team

characteristics and effectiveness: A replication and extension. *Personnel

Psychology, 49*, 429-452. doi://10.1111/j.1744-6570.1996.tb01806.x

Castro, S.L. (2002). Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlation coefficients, $r_{wg(j)}$, hierarchical linear modeling, within- and between-analysis, and random group resampling. *The Leadership Quarterly, 13*, 69-93. doi://10.1016/s1048-9843(01)00105-9

Carroll, V.P., Lee, H.L., & Rao, A.G. (1986). Implications of salesforce productivity heterogeneity and demotivation: a Navy recruiter case study. *Management Science, 32*(11), 1371-1388. doi://10.1287/mnsc.32.11.1371

Chen, G., & Bliese, P.D. (2002). The role of different levels of leadership in predictive self- and collective efficacy: Evidence for discontinuity. *Journal of Applied Psychology, 87*(3), 549-556. doi://10.1037/0021-9010.87.3.549

Clarke, P., & Wheaton, B. (2007). Addressing data sparseness in contextual population research: Using cluster analysis to create synthetic neighborhoods. *Sociological Methods & Research, 35*(3), 311-351. doi://10.1177/0049124106292362

Coglan, D., & Brannick, T. (2005). *Doing action research in your own organization* (2nd Ed). Thousand Oaks, CA: Sage.

Cohen, S.G., & Bailey, D.E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management, 23*(3), 239-290. doi://10.1177/014920639702300303

Combs, J., Liu, Y., Hall, A., & Ketchen. D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology, 59*, 510-528.

Comer, D.R. (1995). A model of social loafing in real work groups. *Human Relations, 48*(6), 647-667. doi://10.1177/001872679504800603

Creswell, J.W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th Edition). Boston, MA: Pearson Education.

Cummings, T.G., & Worley, C.G. (2008). *Organizational development and change* (9th Ed.). Mason, OH: South-Western.

Dansereau, F., Cho, J., & Yammarino, F.J. (2006). Avoiding the "fallacy of the wrong level": A within and between analysis (WABA) approach. *Group & Organization Management, 31*(5), 536-577. doi://10.1177/1059601106291131

de Cremer, D., Snyder M., & DeWitte, S. (2001). 'The less I trust, the less I contribute (or not)?' The effects of trust, accountability, and self-monitoring in social dilemmas. *European Journal of Social Psychology, 31,* 93-107. doi://10.1002/ejsp.34

De Drue, C.K.W. (2010). Social value orientation moderates ingroup love but not outgroup hate in competitive intergroup conflict. *Group Processes & Intergroup Relations, 13*(6), 701-713. doi://10.1177/1368430210377332

De Drue, C.K.W., & Beersma, B. (2005). Conflict in organizations: Beyond effectiveness and performance. *European Journal of Work and Organizational Psychology, 14*(2), 105-117. doi://10.1080/13594320444000227

De Drue, C.K.W., & West, M.A. (2001). Minority dissent and team innovation: The importance of participation in decision making. *Journal of Applied Psychology, 86*(6), 1191-1201. doi://10.1037/0021-9010.86.6.1191

De Jong, A., de Ruyter, K., & Wetzels, M. (2005). Antecedents and consequences of group potency: a study of self-managing service teams. *Management Science, 51*(11), 1610-1625. doi://10.1287/mnsc.1050.0425

Dertouzos, J.N., & Garber, S. (2003). *Is Military Advertising Effective? An Estimation Methodology and Applications to Recruiting in the 1980s and 90s.* Santa Monica, CA: RAND.

Dertouzos, J.N., Polich, J.M., Bamezai, A., & Chesnutt, T. (1989). *Recruiting Effects of Army Advertising*. RAND.

England, G.W. (1975). *The manager and his values: an international perspective from the United States, Japan, Korea, India and Australia.* Ballinger: Cambridge, MA.

Erez, M., & Somech, A. (1996). Is group productivity loss the rule or the exception? Effects of culture and group-based motivation. *Academy of Management Journal, 39*(6), 1513-1537. doi://10.2307/257067

Fernandez-Ballesteros, R., Diez-Nicolas, J., Caprara, G.V., Barbaranelli, C., & Bandura, A. (2002). Determinants and structural relation of personal efficacy to collective efficacy. *Applied Psychology: An International Review, 51*(1), 107-125. doi://10.1111/1464-0597.00081

George, J.M., & James, L.R. (1993). Personality, affect, and behavior in groups revisited: Comment on aggregation, levels of analysis, and a recent application of within and between analysis. *Journal of Applied Psychology, 78*(5), 798-804. doi://10.1037//0021-9010.78.5.798

Gilley, A., & Kerno, S.J. (2010). Groups, teams, and communities of practice: a

comparison. *Advances in Developing Human Resources, 12*(1), 46-60.

doi://10.1177/1523422310365312

Gladstein, D.L. (1984). Groups in context: A model of task group effectiveness.

Administrative Science Quarterly, 29, 499-517. doi://10.2307/2392936

Griesinger, D.W., & Livingston, J.W. (1973). Toward a model of interpersonal

motivation in experimental games. *Behavioral Science, 18,* 173-188.

doi://10.1002/bs.3830180305

Guzzo, R.A., Yost, P.R., Campbell, R.J., & Shea, G.P. (1993). Potency in groups:

Articulating a construct. *British Journal of Social Psychology, 32*, 87-106.

doi://10.1111/j.2044-8309.1993.tb00987.x

Hackman, J.R. (1987). The design of work teams. In Lorsch, J.W. (ed.), *Handbook of*

Organizational Behavior (315-342). Englewood Cliffs, NJ: Prentice-Hall.

Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed

models: the MCMCglmm R package. *Journal of Statistical Software, 33*(2), 1-22.

Halekoh, U. & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric

bootstrap methods for tests in linear mixed models - the R Package pbkrtest.

Journal of Statistical Software, 58(10), 1-30.

Harter, J.K., Schmidt, F.L., & Hayes, T.L. (2002). Business-unit-level relationship

between employee satisfaction, employee engagement, and business outcomes: A

meta-analysis. *Journal of Applied Psychology, 87*(2), 268-279.

doi://10.1037/0021-9010.87.2.268

Headquarters, Department of the Army (2015). *Leader Development* (FM 6-22).

https://armypubs.us.army.mil/epubs/DR_pubs/DR_a/pdf/web/fm6_22.pdf

Headquarters, Department of the Army (2014). *Mission command* (ADP 6-0, C2).

Retrieved from http://armypubs.army.mil/doctrine/DR_pubs/dr_a/pdf/adp6_0.pdf

Hecht, T.D., Allen, N.J., Klammer, J.D., & Kelly, E.C. (2002). Group beliefs, ability, and

performance: The potency of group potency. *Group Dynamics: Theory, Research,*

and Practice, 6(2), 143-152. doi://10.1037//1089-2699.6.2.143

Heppner, P.P., Wampold, B.E., & Kivlighan, D.M. (2008). *Research design in*

counselling (3rd Edition). Belmont, CA: Brooks/Cole.

Heydebrand, W. V. (1973). *Comparative organizations: The results of empirical*

research. Englewood Cliffs, NJ: Prentice-Hall.

Hinsz, V.B., Tindale, R.S., & Vollrath, D.A. (1997). The emerging conceptualization of

groups as information processors. *Psychological Bulletin, 121*(1), 43-64.

doi://10.1037//0033-2909.121.1.43

Hyatt, D.E., & Ruddy, T.M. (1997). An examination of the relationship between work

group characteristics and performance: Once more into the breech. *Personnel*

Psychology, 50, 553-585. doi://10.1111/j.1744-6570.1997.tb00703.x

James, L.R., Demaree, R.G., & Wolf, G. (1984). Estimating within-group interrater

reliability with and without response bias. *Journal of Applied Psychology, 69*(1),

85-98. doi://10.1037//0021-9010.69.1.85

James, L.R., Demaree, R.G., & Wolf, G. (1993). r_{wg}: An assessment of within-group

interrater agreement. *Journal of Applied Psychology, 78*(2), 306-309.

doi://10.1037//0021-9010.78.2.306

Jehn, K A., Chadwick, C., & Thatcher, S. M. (1997). To agree or not to agree: The

 effects of value congruence, individual demographic dissimilarity, and conflict on

 workgroup outcomes. *International Journal of Conflict Management, 8*(4), 287-

 305. doi://10.1108/eb022799

Jex, S.M., & Bliese, P.D. (1999). Efficacy beliefs as a moderator of the impact of work-

 related stressors: a multilevel study. *Journal of Applied Psychology, 84*(3), 349-

 361. doi://10.1037/0021-9010.84.3.349

Joint Committee on Standards for Educational and Psychological Testing of the

 American Education Research Association, American Psychological Research

 Association, and National Council on Measurement in Education. *Standards for*

 Educational and Psychological Testing (2014). Washington: American Education

 Research Council.

Jordan, M.H., Field, H.S., & Armenakis, A.A. (2002). The relationship of group process

 variables and team performance: A team-level analysis in a field setting. *Small*

 Group Research, 33(1), 121-150. doi://10.1177/104649640203300104

Jung, D. I., & Sosik, J. J. (2003). Group potency and collective efficacy: Examining their

 predictive validity, level of analysis, and effects of performance feedback on

 future group performance. *Group & Organization Management, 28*(3), 366-391.

 doi://10.1177/1059601102250821

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from

 restricted maximum likelihood. *Biometrics, 53*, 983-997.

Kleykamp, M.A. (2006). College, jobs, or the military? Enlistment during a time of war. *Social Science Quarterly, 87*(2), 272-290. doi://10.1111/j.1540-6237.2006.00380.x

Knight, G.P. (1981). Behavioral and sociometric methods of identifying cooperators, competitors, and individualists: support for the validity of the social orientation construct. *Developmental Psychology, 17*(4), 430-433. doi://10.1037/0012-1649.17.4.430

Knight, G.P., & Dubro, A.F. (1984). An individualized regression and clustering assessment of the social values of adults and children. *Journal of Research in Personality*, 18, 372-382. doi://10.1016/0092-6566(84)90021-7

Knight, G.P., & Dubro, A.F. (1984). Cooperative, competitive, and individualistic social values: An individualized regression and clustering approach. *Journal of Personality and Social Psychology, 46*(1), 98-105. doi://10.1037/0022-3514.46.1.98

Kolb, J.A. (2016). How not to work and play with others: An updated taxonomy of dysfunctional individual behaviors in groups. *Performance Improvement, 55*(2), 12-20. doi://10.1002/pfi.21553

Kuhlman, D.M., & Marshello, A. (1975). Individual differences in the game motives of own, relative, and joint gain. *Journal of Research in Personality*, 9, 240-251. doi:10.1016/0092-6566(75)90020-3

Liebrand, W.B.G. (1984). The effect of social motives, communication, and group size on behaviour in an N-person multi-stage mixed-motive game. *European Journal of Social Psychology, 14,* 239-264. doi://10.1002/ejsp.2420140302

Liebrand, W.B.G., & McClintock, C.G. (1988). The ring measure of social values: a

    computerized procedure for assessing individual differences in information

    processing and social value orientation. *European Journal of Personality, 2*, 217-

    230. doi://10.1002/per/2410020304

Lindell, M.K., Brandt, C.J., & Whitney, D.J. (1999). A revised index of interrater

    agreement for multi-item ratings of a single target. *Applied Psychological*

    Measurement, 23, 127-135.

Loch, C.H., Huberman, B.A., & Stout, S. (2000). Status competition and performance in

    work groups. *Journal of Economic Behavior & Organization, 43*, 35-55.

    doi://10.1016/s0167-2681(00)00107-4

MacDonald, P. L., & Gardner, R. C. (2000). Type I Error Rate Comparisons of Post Hoc

    Procedures for *I J* Chi-Square Tables. *Educational and Psychological*

    Measurement, 60(5), 735-754. doi://10.1177/00131640021970871

Maki, J.E., & McClintock, C.G. (1983). The accuracy of social value prediction: actor

    and observer influences. *Journal of Personality and Social Psychology, 45*(4),

    829-838. doi://10.1037/0022-3514.45.4.829

Mariadoss, B.J., Milewicz, C., Lee, S., & Sahaym, A. (2014). Salesperson competitive

    intelligence and performance: The role of product knowledge and sales force

    automation usage. *Industrial Marketing Management, 43,* 136-145.

    doi://10.1016/j.indmarman.2013.08.005

Markova, G., & Perry, J.T. (2012). Cohesion and individual well-being of members in

    self-managed teams. *Leadership & Organization Development Journal, 35*(5),

    429-441. doi://10.1108/lodj-04-12-0058

Marks, M.A., Zaccaro, S.J., & Mathieu, J.E. (2000). Performance implications for leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology, 85*(6), 971-986. doi://10.1037/0021-9010.85.6.971

McClintock, C.G., Messick, D.M., Kuhlman, D.M., & Campos, F.T. (1973). Motivational bases of choice in three-choice decomposed games. *Journal of Experimental Social Psychology, 9*, 572-590. doi://10.1016/0022-1031(73)90039-5

McGrath, J.E., & Tschan, F. (2007). Temporal matters in the study of work groups in organizations. *The Psychologist-Manager Journal, 10*(1), 2-12. doi://10.1080/10887150709336609

McIntyre, R.M., & Salas, E. (1995). Measuring and managing for team performance: Lessons from complex environments. In R.A. Guzzo & E. Salas (Eds.), *Team Effectiveness and Decision-Making in Organizations*. San Francisco: Jossey-Bass.

McNeish, D. & Stapleton, L.M. (2014). The effect of small sample size on two-level model estimates: a review and illustration. *Educational Psychological Review, 28*(2), 295-314. doi://10.1007/s10648-014-9287-x

McNeish, D., Stapleton, L.M., & Silverman, R.D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 21*(1), 114-140. doi://10.1037/met0000078

McNeish, D. & Wentzel, K.R. (2016). Accommodating small sample sizes in three level models when the third level is incidental. *Multivariate Behavioral Research, 52*(2), 200-215. doi://10.1080/00273171.2016.1262236

Mehay, S.L., Gue, K.R., & Hogan, P.F. (2000). *The recruiting station location evaluation system (RSLES): A summary report* (NPS Technical Report No. NPS-SM-00-009). Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a382313.pdf.

Messick, D.M., & McClintock, C.G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4,* 1-25. doi://10.1016/0022-1031(68)90046-2

Miles, S.J., & Mangold, G. (2002). The impact of team leader performance on team member satisfaction: the subordinate's perspective. *Team Performance Management: An International Journal, 8*(5/6), 113-121. doi://10.1108/13527590210442230

Mintzberg, H. (1983). The power game and the players. In J.M Shafritz, J.S. Ott & Y.S. Jang (Eds.), *Classics of organizational theory* (8th Ed.). Boston: Cengage.

Moore, L.F., Johns, G., Pinder, C.C. (1980). Toward middle range theory: An overview and perspective. In C.C. Pinder, & Moore, L.F. (Eds.), *Middle Range Theory and the Study of Organizations* (pp.1-16). Hingham MA: Martinus Nijhoff.

Mueller, J.S. (2011). Why individuals in larger teams perform worse. *Organizational Behavior and Human Decision Processes, 117,* 111-124. doi://10.1016/j.obhdp.2011.08.004

Mulvey, P.W., & Klein, H.J. (1998). The impact of perceived loafing and collective efficacy on group goal processes and group performance. *Organizational Behavior and Human Decision Processes, 74*(1), 62-87. doi://10.1006/obhd.1998.2753

Murphy, R.O., & Ackermann, K.A. (2014). Social value orientation: Theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review, 8*(1), 13-41. doi://10.1177/1088868313501745

Newman, D.A., & Sin, H.P. (2009). How do missing data bias estimates of within-group agreement? Sensitivity of SD_{wg}, CV_{wg}, $r_{wg(j)}$, , $r_{wg(j)}$*, and ICC to systematic nonresponse. *Organizational Research Methods, 12*(1), 113-147. doi://10.1177/1094428106298969

O'Connor, K.M., Gruenfeld, D.H., & McGrath, J.E. (1993). The experience and effects of conflict in continuing work groups. *Small Group Research, 24*(3), 362-382. doi://10.1177/1046496493243005

Orvis, B.R., Sastry, N., & McDonald, L.L. (1996). *Military Recruiting Outlook: Recent Trends in Enlistment Propensity and Conversion of Potential Enlisted Supply.* Santa Monica, CA: RAND.

Overbeck, J.R., Correll, J., & Park, B. (2005). Internal status sorting in groups: The problem of too many stars. *Research on Managing Groups and Teams, 7,* 171-202. doi://10.1016/s1534-0856(05)07008-8

Pawitan, Y. (2001). *In all likelihood: Statistical modelling and inference using likelihood.* Oxford, UK: Oxford UP.

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly, 44*(1), 1-28. doi://10.2307/2667029

Piercy, N.F., Cravens, D.W., & Morgan, N.A. (1999). Relationships between sales management control, territory design, salesforce performance and sales

organizational effectiveness. *British Journal of Management, 10*, 95-111.

    doi://10.1111/1467-8551.00113

Posthuma, R.A., Campion, M.C., Masimova, M., & Campion, M.A. (2013). A high

    performance work practices taxonomy: Integrating the literature and directing

    future research. *Journal of Management, 39*, 1184-1220.

R Core Team (2015). R: A language and environment for statistical computing

    [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.

Raes, E., Kyndt, E., Decuyper, S., Van Den Bossche, P., & Dochy, F. (2015). An

    exploratory study of group development and team learning. *Human Resource*

    Development Quarterly, 26(1), 5-30. doi://10.1002/hrdq.21201

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and*

    data analysis methods (2nd ed.). Thousand Oaks, CA: Sage.

Rego, A., Junior, D.R., & Cunha, M.P. (2015). Authentic leaders promoting store

    performance: The mediating roles of virtuousness and potency. *Journal of*

    Business Ethics, 128, 617-634. doi://10.1007/s10551-014-2125-8

Russ-Eft, D. (2016). Controversies that shaped the field of human resource development:

    Town hall forums of the Academy of Human Resource Development. *Advances*

    in Developing Human Resources, 18(4), 512-535.

    doi://10.1177/1523422316660148

Sakalaki, M., & Sotiriou, P. (2012). Pro-self orientation and preference for deceitful

    strategies: Social value orientation, dispositional and behavioral correlates of

    economic opportunism. *Studia Psychologica, 54*(2), 157-165.

Scott-Ladd, B., Travaglione, A., & Marshall, V. (2006). Casual inferences between participation in decision-making, task attributes, work effort, rewards, job satisfaction, and commitment. *Leadership & Organizational Development Journal, 27*(5), 399-414. doi://10.1108/01437730610677990

Segal, D.R., Bachman, J.G., Freedman-Doan, P., & O'Malley, P.M. (1999). Propensity to serve in the U.S. military: Temporal trends and subgroup differences. *Armed Forces & Society, 25*(3), 407-427. doi://10.1177/0095327x9902500304

Shavelson, R.J., Haggstrom, G.W. & Winkler, J.D. (1983) *Potential for Military Recruiting from Two-Year Colleges and Postsecondary Vocational Schools.* Santa Monica, CA: RAND.

Shieh, G. (2012). A comparison of two indices for the intraclass correlation. *Behavioral Research, 44,* 1212-1223. doi://10.3758/s13428-012-0188-y

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428. doi://10.1037//0033-2909.86.2.420

Siders, M.A., George, G., & Dharwadkar, R. (2001). The relationship of internal and external commitment foci to objective job performance measures. *Academy of Management Journal, 44*(3), 570-579. doi://10.2307/3069371

Smith-Crowe, K., Burke, M. J., Cohen, A., & Doveh, E. (2014). Statistical significance criteria for the r_{wg} and average deviation interrater agreement indices. *Journal of Applied Psychology, 99*, 239-261.

Sonnemans, J., van Dijk, F., & van Winden, F.A.A.M. (2001). *On the dynamics of social ties structures in groups.* Amsterdam: CrEED, Department of Economics.

Retrieved from http://dare.uva.nl/search?identifier=16f551e8-d20c-4a29-b28f-8f5fba09a864

Stajkovic, A.D., Lee, D., & Nyberg, A.J. (2009). Collective efficacy, group potency, and group performance: meta-analyses of their relationships, and test of a mediation model. *Journal of Applied Psychology, 94*(3), 811-828. doi://10.1037/a0015659

Stevens, M.J., & Campion, M.A. (1994). The knowledge, skills, and ability requirements for teamwork: Implications for human resource management. *Journal of Management, 20*(2), 503-530. doi://10.1177/014920639402000210

Stouten, J., de Cremer, D., & van Dijk, E. (2005). All is well that ends well, at least for proselfs: Emotional reactions to equality violations as a function of social value orientation. *European Journal of Social Psychology, 35,* 767-783. doi://10.1002/ejsp.276

Sundstrom, E., De Meuse, K.P., & Futrell, D. (1990). Work teams: applications and effectiveness. *American Psychologist, 42*(2), 120-133. doi://10.1037/0003-066X.45.2.120

Sutterlin, B., Brunner, T.A., & Siegrist, M. (2013). Impact of social value orientation on energy conservation in different behavioral domains. *Journal of Applied Social Psychology, 43,* 1725-1735. doi://10.1111/jasp.12128

Swanson, R.A., & Holton III, E.F. (2009). *Foundations of human resource development* (2nd ed.). San Francisco: Barrett-Koehler.

Tarricone, P., & Luca, J. (2002). Employees, teamwork and social interdependence – formula for successful business? *Team Performance Management: An International Journal, 8*(3/4), 54-59. doi://10.1108/13527590210433348

United States Department of Defense (2012). *2012 Demographics: Profile of the Military Community.*

http://download.militaryonesource.mil/12038/MOS/Reports/2012_Demographics_Report.pdf

Upton, D.R. (2009). Implications of social value orientation and budget levels on group performance and performance variance. *Journal of Management Accounting Research, 21*, 293-316. doi://10.2308/jmar.2009.21.1.293

Van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S.A.R.B., & Crone, E.A. (2009). What motivates repayment? Neutral correlates of reciprocity in the trust game. *Social Cognitive & Affective Neuroscience*, 4, 294-304. doi://10.1093/scan/nsp009

Van Lange, P.A.M., & Leibrand, W.B.G. (1991). The influence of other's morality and own social value orientation on cooperation in The Netherlands and the U.S.A. *International Journal of Psychology, 26*(4), 429-449. doi://10.1037/0022-3514.73.4.733

Van Lange, P.A.M., Otten, E., De Bruin, E.M.N., & Joireman, J. (1997). Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *Journal of Personality and Social Psychology, 73*(4), 733-746. doi://10.1037/0022-3514.73.3.733

van Vugt, M., Jepson, S., Hartman, C., & de Cremer, D. (2004). Why autocratic leadership might fail in solving public good dilemmas. *Journal of Experimental Psychology, 40,* 1-20.

Venneberg, D.L. (2010). An eclectic approach to building effective teams: eight separate

    but interrelated components. *Advances in Developing Human Resources, 12*(1),

    3-6. doi://10.1177/1523422310365719

Verint (2014). Verint Enterprise Feedback Management [Computer software].

    Information retrieved from

    http://www.verint.com/Assets/verint/resources/datasheets/EFM.pdf.

Wagenmakers, E.J., & Farrell, S. (2004). AIC model selection using Akaike weights.

    Psychonomic Bulletin & Review, 11(1), 192-196.

Wagner III, J.A. (1995). Studies of individual-collectivism: Effects on cooperation in

    groups. *Academy of Management Journal, 38*(1), 152-172. doi://10.2307/256731

Wheelan, S.A. (2009). Group size, group development, and group productivity. *Small

    Group Research, 40*(2), 247-262. doi://10.1177/1046496408328703

Williams, B.C., & Plouffe, C.R. (2007). Assessing the evolution of sales knowledge: A

    20-year content analysis. *Industrial Marketing Management, 36*, 408-419.

    doi://10.1016/j.indmarman.2005.11.003

Whitely, W. (1979). A cross-national test of England's model of manager's value

    systems and their relationship to behavior. In G.W. England, A.R. Negandhi, &

    B. Wilpert (Eds.), *Organizational function in a cross-cultural perspective*. Kent

    State, OH: Comparative Administration Research Institute.

Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A.,

    Inukai, K., Tanida, S., Kiyonari, T., Takagishi, H., & Simunovic, D. (2013). Is

    behavioral pro-sociality game-specific? Pro-social preference and expectations of

pro-sociality. *Organizational Behavior and Human Decision Processes, 120*, 260-271. doi://10.1016/j.obhdp.2012.06.002

Zampetakis, L.A. (2014). Sales force management practices in organizations with a supportive climate towards creativity. *Journal of Strategic Marketing, 22*(1), 59-72.

Zander, A. (1985). *The purposes of groups and organizations.* San Francisco: Jossey-Bass.

Appendix A

Instrument

Notes:
-Subscale titles, references, and comments are omitted from the survey (i.e. while they are annotated here for reference, they were not included in the live survey to Army Recruiters)
-Respondents are not required to answer any question; they will be able to "click through" any page to the next without responding to one if they wish
-Item order was randomized within subscale

Group Potency
Adapted from Guzzo, Yost, Campbell & Shea (1993). Minor edits have been made to adapt these items to a recruiting center (i.e. replaced "my work group" with "my Recruiting Center", etc.)
Likert-type scale (5-pt, strongly disagree / disagree / neither agree nor disagree / agree / strongly agree)
1. My Recruiting Center has confidence in itself.
2. My Recruiting Center believes it can become unusually good at performing its job.
3. My Recruiting Center expects to be known as a high-performing unit.
4. My Recruiting Center feels it can solve any problem it encounters.
5. My Recruiting Center believes it can be very productive.
6. My Recruiting Center can get a lot done when it works hard.
7. No task is too tough for my Recruiting Center.
8. My Recruiting Center expects to have a lot of influence within the Company.

Self-Management
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center (i.e. replaced "my manager" with "our Center Leader")
Likert-type scale (5-pts)
9. The members of my Recruiting Center are responsible for determining the methods, procedures, and schedules with which the work gets done
10. My Recruiting Center rather than our Center Leader decides who does what tasks within the team
11. Most work-related decisions are made by the members of my Recruiting Center rather than by our Center Leader

Participation
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center.
Likert-type scale (5-pts)
12. As a member of a Recruiting Center, I have real say in how the team carries out its work
13. Most members of my Recruiting Center get a chance to participate in decision making

139

14. My Recruiting Center is designed to let everyone participate in decision making

Task Variety
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center.
Likert-type scale (5-pts)
15. Most members of my Recruiting Center get a chance to learn the different tasks the team performs
16. Most everyone in my Recruiting Center gets a chance to do the more interesting tasks
17. Task assignments often change from day to day to meet the work load needs of the Recruiting Center

Task Significance
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center (changed "the customers in our area" to "the people in our area"; changed "contribution to serving the company's customers" to "contribution to the Army"; changed "my work is important to the company" to "my work is important")
Likert-type scale (5-pts)
18. The work performed by my Recruiting Center is important to the people in our area
19. My Recruiting Center makes an important contribution to the Army
20. My Recruiting Center helps me feel that my work is important

Social Value Orientation
Adapted from Van Lange, Otten, De Bruin & Joireman (1997)

Instructions: We ask you to imagine that you have been paired randomly with another person, whom we will refer to simply as the "other". This person is someone you do not know and that you will not knowingly meet in the future. Both you and the "other" person will be making choices by selecting the letter A, B, or C. Your own choices will produce points for both you and the "other" person. Likewise, the other's choices will produce points for him/her and for you. Every point has value: the more points you receive, the better for you, and the more points the "other" receives, the better for him/her.

Before making choices, please keep in mind that there are no right or wrong answers – choose the option that you, for whatever reason, prefer most. Also, remember that the points have value: the more of them you accumulate, the better for you. Likewise, from the "other's" point of view, the more points s/he accumulates, the better for him/her.

		A	B	C
21.	You get	480	540	480
	Other gets	80	280	480
22.	You get	560	500	500
	Other gets	300	500	100
23.	You get	520	520	580

	Other gets	520	120	320
24.	You get	500	560	490
	Other gets	100	300	490
25.	You get	560	500	490
	Other gets	300	500	90
26.	You get	500	500	570
	Other gets	500	100	300
27.	You get	510	560	510
	Other gets	510	300	110
28.	You get	550	500	500
	Other gets	300	100	500
29.	You get	480	490	540
	Other gets	100	490	300

Interdependence
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center (such as the list of job-related rewards in 38, etc.)
Likert-type scale (5-pts)
30. I cannot accomplish my tasks without information or materials from other members of my Recruiting Center
31. Other members of my Recruiting Center depend on me for information or materials needed to perform their tasks
32. Within my Recruiting Center, jobs performed by Recruiters are related to one another
33. My work goals come directly from the goals of my Recruiting Center
34. My work activities on any given day are determined by my Recruiting Center's goals for that day
35. I do very few activities on my job that are not related to the goals of my Recruiting Center
36. Feedback about how well I am doing my job comes primarily from information about how well the entire Recruiting Center is doing
37. My performance evaluation is strongly influenced by how well my Recruiting Center performs
38. Many rewards from my job (e.g. promotion, evaluations, selection for schools, etc.) are determined in large part by my contributions as a member of my Recruiting Center

Relative Size
Adapted from Campion, Medsker & Higgs (1993). Significant edit (changed from an assessment of too-smallness to an assessment of right-sizedness).
Likert-type scale (5-pts):

a. the center would be most effective with more than 2 more required recruiters (i.e. it's much too small)

b. the center would be most effective with 1 or 2 more required recruiters (i.e. it's a bit too small)

c. the center would be most effective with the current number of required recruiters (i.e. it's about right)

d. the center would be most effective with 1 or 2 fewer required recruiters (i.e. it's a bit too large)

e. the center would be most effective with more than 2 fewer required recruiters (i.e. it's much too large)

39. In terms of the number of people (required recruiters; not currently-assigned recruiters) at your Recruiting Center that would seem to you to be the most productive, your assessment is that the current number of required recruiters (including the center leader) is:

Training
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center (specifically changes "the company provides… for my team" to "My Recruiting Center is provided…"; also changed "quality and customer service" to "interacting with applicants and potential applicants")
Likert-type scale (5-pts)

40. My Recruiting Center is provided adequate technical training

41. My Recruiting Center is provided adequate training about interacting with applicants and potential applicants

42. My Recruiting Center is provided adequate team skills training (e.g. communication, organization, interpersonal skills, etc.)

Social Support
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center
Likert-type scale (5-pts)

43. Being in my Recruiting Center gives me the opportunity to work in a team and provide support to other team members.

44. My Recruiting Center increases my opportunities for positive social interaction.

45. Members of my Recruiting Center help each other out at work when needed.

Workload Sharing
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center (i.e. replaced "my manager" with "our Center Leader")
Likert-type scale (5-pts)

46. Everyone in my Recruiting Center does their fair share of the work.

47. No one in my Recruiting Center depends on other team members to do the work for them.

48. Nearly all the members of my Recruiting Center contribute equally to the work.

Communication / Cooperation Within the Work Group
Adapted from Campion, Medsker & Higgs (1993). Minor edits have been made to adapt these items to a recruiting center (i.e. replaced "my manager" with "our Center Leader")
Likert-type scale (5-pts)
49. Members of my Recruiting Center are very willing to share information with each other about our work.
50. My Recruiting Center enhances the communication amongst its recruiters.
51. Members of my Recruiting Center cooperate to get the work done.

Appendix B

R Syntax and Output for Null Model Calculations

Null Model: Productivity
##Calculate Null (unconditional) model for each outcome variable (production per recruiter and mission accomplishment)
nullmodel_PPRRF<-lmer(PPRRF ~ 1 + (1 | RSID3), data = grp_vars)
summary(nullmodel_PPRRF)
>   Linear mixed model fit by REML ['lmerMod']
>   Formula: PPRRF ~ 1 + (1 | RSID3)
>   Data: grp_vars
>
>   REML criterion at convergence: -58.8
>
>   Scaled residuals:
>   Min   1Q Median   3Q   Max
>   -2.1260 -0.6106 -0.1955 0.5169 2.8561
>
>   Random effects:
>   Groups   Name   Variance Std.Dev.
>   RSID3   (Intercept) 0.01142 0.1069
>   Residual   0.03269 0.1808
>   Number of obs: 179, groups: RSID3, 54
>
>   Fixed effects:
>   Estimate Std. Error t value
>   (Intercept)   0.6603   0.0202   32.69

Null Model: Managerial Judgement
nullmodel_MA<-lmer(MA_PERC ~ 1 + (1 | RSID3), data = grp_vars)
summary(nullmodel_MA)
>   Linear mixed model fit by REML ['lmerMod']
>   Formula: MA_PERC ~ 1 + (1 | RSID3)
>   Data: grp_vars
>
>   REML criterion at convergence: 1569.9
>
>   Scaled residuals:
>   Min   1Q Median   3Q   Max
>   -2.46694 -0.67032 0.08989 0.65722 2.75887
>
>   Random effects:
>   Groups   Name   Variance Std.Dev.
>   RSID3   (Intercept) 29.32   5.414
>   Residual   358.71   18.940

Number of obs: 179, groups: RSID3, 54

Fixed effects:
          Estimate Std. Error t value
(Intercept)   86.889   1.614   53.83

Appendix C

R Syntax and Output for Hypothesis 1 Calculations

Grand Model: Productivity

##PRODUCTIVITY PER REQUIRED RECRUITER HYPOTHESIS 1
grand_H1_model_PPRRF<-lmer(PPRRF ~ U1_S + RRF_S + QMA_S + PROP_S + (1 |
RSID3), data = grp_vars)
summary(grand_H1_model_PPRRF)
    Linear mixed model fit by REML
    t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
    Formula: PPRRF ~ U1_S + RRF_S + QMA_S + PROP_S + (1 | RSID3)
      Data: grp_vars

    REML criterion at convergence: -35

    Scaled residuals:
        Min   1Q   Median   3Q   Max
    -2.23745 -0.58436 -0.09874 0.51066 3.15545

    Random effects:
     Groups   Name   Variance Std.Dev.
     RSID3   (Intercept) 0.009913 0.09957
     Residual   0.032405 0.18001
    Number of obs: 179, groups: RSID3, 54

    Fixed effects:
                  Estimate Std. Error   df t value Pr(>|t|)
    (Intercept) 6.752e-01 2.746e-02 4.666e+01 24.590   <2e-16 ***
    U1_S   2.531e-02 1.917e-02 5.107e+01   1.320   0.193
    RRF_S   -1.007e-02 8.353e-03 1.732e+02 -1.206   0.229
    QMA_S   -4.062e-05 1.612e-03 1.702e+02 -0.025   0.980
    PROP_S   1.417e-02 1.075e-02 5.243e+01   1.318   0.193

    Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    Correlation of Fixed Effects:
          (Intr) U1_S   RRF_S QMA_S
    U1_S -0.338
    RRF_S 0.058 0.016
    QMA_S -0.082 0.132 -0.744
    PROP_S 0.679 -0.253 -0.050 0.040

Limited Model: Productivity

limited_H1_model_PPRRF<-lmer(PPRRF ~ U1_S + RRF_S + (1 | RSID3), data = grp_vars)
summary(limited_H1_model_PPRRF)
      Linear mixed model fit by REML
      t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
      Formula: PPRRF ~ U1_S + RRF_S + (1 | RSID3)
        Data: grp_vars

      REML criterion at convergence: -51.5

      Scaled residuals:
        Min   1Q Median   3Q   Max
      -2.1180 -0.6076 -0.0886 0.5214 3.1838

      Random effects:
       Groups   Name   Variance Std.Dev.
       RSID3   (Intercept) 0.009908 0.09954
       Residual   0.032352 0.17987
      Number of obs: 179, groups: RSID3, 54

      Fixed effects:
            Estimate Std. Error   df t value Pr(>|t|)
      (Intercept) 0.650409 0.019928 46.530000 32.637 <2e-16 ***
      U1_S   0.031909 0.018337 49.500000 1.740 0.0880 .
      RRF_S   -0.010003 0.005575 171.200000 -1.794 0.0745 .

      Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      Correlation of Fixed Effects:
         (Intr) U1_S
      U1_S -0.217
      RRF_S 0.022 0.170

Model Comparisons: Prodictivity
> KRmodcomp(grand_H1_model_PPRRF,nullmodel_PPRRF)
F-test with Kenward-Roger approximation; computing time: 0.11 sec.
large : PPRRF ~ U1_S + RRF_S + QMA_S + PROP_S + (1 | RSID3)
small : PPRRF ~ 1 + (1 | RSID3)
     stat   ndf   ddf F.scaling p.value
Ftest 2.2851 4.0000 96.6215   0.99542 0.06569 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> KRmodcomp(limited_H1_model_PPRRF,nullmodel_PPRRF)

F-test with Kenward-Roger approximation; computing time: 0.06 sec.
large : PPRRF ~ U1_S + RRF_S + (1 | RSID3)
small : PPRRF ~ 1 + (1 | RSID3)
     stat ndf   ddf F.scaling p.value
Ftest 3.711 2.000 93.351   0.99574 0.02813 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> KRmodcomp(grand_H1_model_PPRRF,limited_H1_model_PPRRF)
F-test with Kenward-Roger approximation; computing time: 0.05 sec.
large : PPRRF ~ U1_S + RRF_S + QMA_S + PROP_S + (1 | RSID3)
small : PPRRF ~ U1_S + RRF_S + (1 | RSID3)
      stat   ndf   ddf F.scaling p.value
Ftest 0.8663 2.0000 93.0384   0.9975 0.4239

Grand Model: Managerial Judgement

> grand_H1_model_MA<-lmer(MA_PERC ~ U1_S + RRF_S + QMA_S + PROP_S + (1 | RSID3), data = grp_vars)
> summary(grand_H1_model_MA)
     Linear mixed model fit by REML
     t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
     Formula: MA_PERC ~ U1_S + RRF_S + QMA_S + PROP_S + (1 | RSID3)
      Data: grp_vars

     REML criterion at convergence: 1562.7

     Scaled residuals:
       Min   1Q Median   3Q   Max
    -2.3536 -0.7565 0.0161 0.5973 2.7160

     Random effects:
     Groups   Name   Variance Std.Dev.
     RSID3   (Intercept) 23.1   4.807
     Residual   363.2   19.059
     Number of obs: 179, groups: RSID3, 54

     Fixed effects:
           Estimate Std. Error   df t value Pr(>|t|)
     (Intercept) 85.1769   2.2513 41.5700 37.835 <2e-16 ***
     U1_S   2.5158   1.5892 45.1900 1.583   0.120
     RRF_S   0.8978   0.8151 159.5800 1.101   0.272
     QMA_S   -0.1545   0.1527 122.4100 -1.012   0.314
     PROP_S   -0.8741   0.8977 50.5700 -0.974   0.335

     Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
```
        (Intr) U1_S   RRF_S  QMA_S
U1_S   -0.353
RRF_S   0.082  0.034
QMA_S  -0.107  0.144 -0.734
PROP_S  0.676 -0.268 -0.060  0.050
```

Model Comparisons: Managerial Judgement

> KRmodcomp(grand_H1_model_MA,nullmodel_MA)
```
        F-test with Kenward-Roger approximation; computing time: 0.06 sec.
        large : MA_PERC ~ U1_S + RRF_S + QMA_S + PROP_S + (1 | RSID3)
        small : MA_PERC ~ 1 + (1 | RSID3)
                stat    ndf    ddf F.scaling p.value
        Ftest  1.0578  4.0000 85.9736   0.99689  0.3825
```

Secondary Estimations

```
#MCMC
library(MCMCglmm)
#to use interaction terms use variable:variable
H1_PPRRF_MCMC<-MCMCglmm(PPRRF<-PPRRF ~ U1_S + RRF_S, random = ~
RSID3, nitt=30000, burnin = 2000, thin = 1, data = grp_vars)
summary(H1_PPRRF_MCMC)
        Iterations = 2001:30000
        Thinning interval  = 1
        Sample size  = 28000

        DIC: -72.33603

        G-structure:  ~RSID3

            post.mean  l-95% CI u-95% CI eff.samp
        RSID3  0.008998 2.006e-08  0.01769    862.2

        R-structure:  ~units

            post.mean l-95% CI u-95% CI eff.samp
        units   0.03397  0.02522  0.04368     1007

        Location effects: PPRRF ~ U1_S + RRF_S

                post.mean   l-95% CI   u-95% CI eff.samp  pMCMC
        (Intercept)  0.6505887  0.6116960  0.6892173    28000 <4e-05 ***
        U1_S        0.0311816 -0.0044788  0.0669738    24605 0.0856 .
```

RRF_S   -0.0101516 -0.0213340 0.0008356   28000 0.0744 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> H2_MA_MCMC<-MCMCglmm(MA_PERC<-MA_PERC ~ POTENCY_S + RELATI
VE_SIZE_S + COOPERATION_S, random = ~ RSID3, nitt=30000, burnin = 2000, thin
= 1, data = grp_vars)
> summary(H2_MA_MCMC)

Iterations = 2001:30000
Thinning interval = 1
Sample size = 28000

DIC: 1531.022

G-structure: ~RSID3

      post.mean l-95% CI u-95% CI eff.samp
RSID3   3.09 1.769e-09   19.58   187.3

R-structure: ~units

      post.mean l-95% CI u-95% CI eff.samp
units   294   231.8   356.9   7799

Location effects: MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPE
RATION_S

              post.mean l-95% CI u-95% CI eff.samp   pMCMC
(Intercept)   86.8102 84.2171 89.3599   28000 < 4e-05 ***
POTENCY_S   23.2386 16.8791 29.6987   28000 < 4e-05 ***
RELATIVE_SIZE_S   4.0919 -0.2054   8.5179   28000 0.06571 .
COOPERATION_S   -9.0662 -14.6172 -3.5598   29334 0.00121 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> H1_MA_MCMC<-MCMCglmm(MA_PERC<-MA_PERC ~ 1, random = ~ RSID3, nit
t=30000, burnin = 2000, thin = 1, data = grp_vars)
> summary(H1_MA_MCMC)

Iterations = 2001:30000
Thinning interval = 1
Sample size = 28000

DIC: 1577.45

G-structure: ~RSID3

     post.mean l-95% CI u-95% CI eff.samp
RSID3   1.713 2.732e-26   8.483   137.7

 R-structure: ~units

    post.mean l-95% CI u-95% CI eff.samp
units   389.5   310.2   475.1   6923

 Location effects: MA_PERC ~ 1

         post.mean l-95% CI u-95% CI eff.samp pMCMC
(Intercept)   86.81   83.94   89.74   28000 <4e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix D

R Syntax and Output for Hypothesis 2 Calculations

Grand Model: Productivity
> grand_H2_model_PPRRF<-lmer(PPRRF ~ U1_S + RRF_S + POTENCY_S + SELF_
MANAGE_S + PARTICIPATION_S + TASK_V_S + TASK_S_S + INTERDEPENDE
NCE_S + RELATIVE_SIZE_S + TRAINING_S + WORKLOAD_SHARING_S + SOCI
AL_SUPPORT_S + COOPERATION_S + (1 | RSID3), data = grp_vars)
> summary(grand_H2_model_PPRRF)
        Linear mixed model fit by REML
        t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
        Formula: PPRRF ~ U1_S + RRF_S + POTENCY_S + SELF_MANAGE_S + PA
        RTICIPATION_S +
            TASK_V_S + TASK_S_S + INTERDEPENDENCE_S + RELATIVE_SIZE_S
        +
            TRAINING_S + WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S + CO
        OPERATION_S +   (1 | RSID3)
          Data: grp_vars

        REML criterion at convergence: -31.2

        Scaled residuals:
            Min   1Q Median   3Q   Max
        -1.9778 -0.6179 -0.0814 0.4930 3.4283

        Random effects:
         Groups   Name   Variance Std.Dev.
         RSID3   (Intercept) 0.005038 0.07098
         Residual   0.031226 0.17671
        Number of obs: 179, groups: RSID3, 54

        Fixed effects:
                        Estimate Std. Error   df t value Pr(>|t|)
        (Intercept)   0.653931 0.017110 38.510000 38.218 < 2e-16 ***
        U1_S   0.017217 0.016300 41.930000   1.056 0.29690
        RRF_S   -0.009333 0.005866 165.000000 -1.591 0.11353
        POTENCY_S   0.174433 0.053016 154.410000   3.290 0.00124 **
        SELF_MANAGE_S   -0.003106 0.035831 162.540000 -0.087 0.93103
        PARTICIPATION_S   0.003720 0.047761 164.450000   0.078 0.93801
        TASK_V_S   -0.005210 0.065032 164.010000 -0.080 0.93624
        TASK_S_S   0.005601 0.055938 151.300000   0.100 0.92038
        INTERDEPENDENCE_S -0.008107 0.054062 153.560000 -0.150 0.88100
        RELATIVE_SIZE_S   0.016084 0.025030 162.280000   0.643 0.52141
        TRAINING_S   0.038901 0.040393 152.900000   0.963 0.33703
        WORKLOAD_SHARING_S -0.014591 0.030248 163.140000 -0.482 0.63018

SOCIAL_SUPPORT_S   0.050359   0.076686 161.550000   0.657 0.51231
COOPERATION_S   -0.129106   0.066663 164.970000 -1.937 0.05449 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Limited Model: Productivity

> limited_H2_model_PPRRF<-lmer(PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + (1 | RSID3), data = grp_vars)
> summary(limited_H2_model_PPRRF)
Linear mixed model fit by REML
t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + (1 | RSID3)
   Data: grp_vars

REML criterion at convergence: -69.3

Scaled residuals:
    Min   1Q Median   3Q   Max
-2.0193 -0.6726 -0.1176 0.5038 3.4464

Random effects:
 Groups   Name   Variance Std.Dev.
 RSID3   (Intercept) 0.005734 0.07572
 Residual   0.029517 0.17181
Number of obs: 179, groups: RSID3, 54

Fixed effects:
               Estimate Std. Error   df t value Pr(>|t|)
(Intercept)   0.653451   0.017193 39.770000 38.007 < 2e-16 ***
U1_S   0.018765   0.016007 42.940000   1.172 0.24752
RRF_S   -0.010013   0.005329 172.200000 -1.879 0.06196 .
POTENCY_S   0.192997   0.035369 170.920000   5.457 1.68e-07 ***
COOPERATION_S -0.079854   0.030460 168.380000 -2.622 0.00955 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) U1_S   RRF_S POTENC
U1_S   -0.218
RRF_S   0.031 0.181
POTENCY_S   0.024 -0.135 -0.036
COOPERATION 0.003 0.081 0.187 -0.647

Model Comparisons: Productivity

> KRmodcomp(grand_H2_model_PPRRF,limited_H1_model_PPRRF)
F-test with Kenward-Roger approximation; computing time: 0.11 sec.
large : PPRRF ~ U1_S + RRF_S + POTENCY_S + SELF_MANAGE_S + PART
ICIPATION_S +
   TASK_V_S + TASK_S_S + INTERDEPENDENCE_S + RELATIVE_SIZE_S
+
   TRAINING_S + WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S + CO
OPERATION_S +
   (1 | RSID3)
small : PPRRF ~ U1_S + RRF_S + (1 | RSID3)
        stat   ndf   ddf F.scaling p.value
Ftest 2.9697 11.0000 159.4264   0.99954 0.001304 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> KRmodcomp(limited_H2_model_PPRRF,limited_H1_model_PPRRF)
F-test with Kenward-Roger approximation; computing time: 0.06 sec.
large : PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + (1 | RS
ID3)
small : PPRRF ~ U1_S + RRF_S + (1 | RSID3)
        stat   ndf   ddf F.scaling   p.value
Ftest 15.264   2.000 171.477   0.99999 7.932e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> KRmodcomp(grand_H2_model_PPRRF,limited_H2_model_PPRRF)
F-test with Kenward-Roger approximation; computing time: 0.06 sec.
large : PPRRF ~ U1_S + RRF_S + POTENCY_S + SELF_MANAGE_S + PART
ICIPATION_S +
   TASK_V_S + TASK_S_S + INTERDEPENDENCE_S + RELATIVE_SIZE_S
+
   TRAINING_S + WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S + CO
OPERATION_S +
   (1 | RSID3)
small : PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + (1 | R
SID3)
        stat   ndf   ddf F.scaling p.value
Ftest 0.3027 9.0000 159.3289   0.99966   0.973

Grand Model: Managerial Judgement

> grand_H2_model_MA<-lmer(MA_PERC ~ POTENCY_S + SELF_MANAGE_S + PA
RTICIPATION_S + TASK_V_S + TASK_S_S + INTERDEPENDENCE_S + RELATI
VE_SIZE_S + TRAINING_S + WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S
+ COOPERATION_S + (1 | RSID3), data = grp_vars)
> summary(grand_H2_model_MA)
Linear mixed model fit by REML

t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: MA_PERC ~ POTENCY_S + SELF_MANAGE_S + PARTICIPATIO
N_S + TASK_V_S +
  TASK_S_S + INTERDEPENDENCE_S + RELATIVE_SIZE_S + TRAINING
_S +
  WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S + COOPERATION_S
+ (1 | RSID3)
  Data: grp_vars

REML criterion at convergence: 1467.7

Scaled residuals:
    Min   1Q Median   3Q   Max
-2.5928 -0.6474 0.1692 0.6569 2.7989

Random effects:
 Groups   Name   Variance Std.Dev.
 RSID3   (Intercept) 8.74   2.956
 Residual   294.05   17.148
Number of obs: 179, groups: RSID3, 54

Fixed effects:
               Estimate Std. Error   df t value Pr(>|t|)
(Intercept)   86.8350   1.3509 41.4800 64.280 < 2e-16 ***
POTENCY_S   22.3607   4.8881 164.5100   4.574 9.35e-06 ***
SELF_MANAGE_S   2.6829   3.3224 166.9000   0.808 0.4205
PARTICIPATION_S -2.9001   4.3887 164.2900 -0.661 0.5097
TASK_V_S   2.5288   5.9877 166.2700   0.422 0.6733
TASK_S_S   1.2781   5.2320 163.2800   0.244 0.8073
INTERDEPENDENCE_S -4.0082   5.0082 160.7100 -0.800 0.4247
RELATIVE_SIZE_S   4.1325   2.2977 166.1600   1.798 0.0739 .
TRAINING_S   1.4320   3.7805 165.7900   0.379 0.7053
WORKLOAD_SHARING_S -2.9040   2.7094 167.0000 -1.072 0.2853
SOCIAL_SUPPORT_S -0.5362   7.1193 166.4800 -0.075 0.9401
COOPERATION_S   -6.9731   6.0714 165.9600 -1.149 0.2524

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
        (Intr) POTENC SELF_M PARTIC TASK_V TASK_S INTERD RELATI
TRAINI WORKLO SOCIAL
POTENCY_S   -0.001
SELF_MANAGE -0.005 -0.110
PARTICIPATI 0.004 0.000 -0.606
TASK_V_S   -0.004 0.022 0.300 -0.553
TASK_S_S   0.003 -0.598 -0.039 0.049 -0.335

INTERDEPEND -0.001 -0.086 -0.137 0.111 -0.247 0.009
RELATIVE_SI 0.001 0.043 -0.081 0.058 -0.054 -0.006 0.012
TRAINING_S -0.002 -0.113 -0.003 0.018 -0.275 0.042 -0.053 0.085
WORKLOAD_SH 0.000 -0.082 0.052 -0.091 -0.016 -0.004 0.056 -0.078 -0.007
SOCIAL_SUPP -0.001 0.086 -0.059 -0.139 0.152 -0.343 0.095 -0.162 -0.222 -0
.011
COOPERATION 0.003 -0.095 0.062 0.050 -0.176 0.236 -0.166 0.097 -0.091 -
0.306 -0.716

Model Comparisons: Managerial Judgement
> KRmodcomp(grand_H2_model_MA,limited_H2_model_MA)
F-test with Kenward-Roger approximation; computing time: 0.06 sec.
large : MA_PERC ~ POTENCY_S + SELF_MANAGE_S + TASK_V_S + TAS
K_S_S + INTERDEPENDENCE_S +
  RELATIVE_SIZE_S + TRAINING_S + WORKLOAD_SHARING_S + SOCI
AL_SUPPORT_S +
  COOPERATION_S + (1 | RSID3)
small : MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION_
S + (1 |
  RSID3)
       stat   ndf   ddf F.scaling p.value
Ftest 0.3022 7.0000 165.3596 0.99968 0.9521

> KRmodcomp(limited_H2_model_MA,nullmodel_MA)
F-test with Kenward-Roger approximation; computing time: 0.07 sec.
large : MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION_S
+ (1 |
  RSID3)
small : MA_PERC ~ 1 + (1 | RSID3)
       stat   ndf   ddf F.scaling p.value
Ftest 19.219 3.000 173.940 0.99998 8.24e-11 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Secondary Estimations

> H2_PPRRF_MCMC<-MCMCglmm(PPRRF<-PPRRF ~ U1_S + RRF_S + POTENCY
_S + COOPERATION_S, random = ~ RSID3, nitt=30000, burnin = 2000, thin = 1, data
= grp_vars)
> summary(H2_PPRRF_MCMC)

   Iterations = 2001:30000
   Thinning interval = 1
   Sample size = 28000

   DIC: -89.77499

G-structure: ~RSID3

```
      post.mean l-95% CI u-95% CI eff.samp
RSID3  0.002921 9.31e-16  0.01011   216.5
```

R-structure: ~units

```
     post.mean l-95% CI u-95% CI eff.samp
units   0.03264   0.0243  0.04153   496.5
```

Location effects: PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S

```
            post.mean   l-95% CI   u-95% CI eff.samp  pMCMC
(Intercept)    0.6537643  0.6220371  0.6845241    28000 <4e-05 ***
U1_S           0.0166052 -0.0133642  0.0452960    12678 0.2564
RRF_S         -0.0098874 -0.0204176  0.0006889    28000 0.0654 .
POTENCY_S      0.2014774  0.1303787  0.2737624     5620 <4e-05 ***
COOPERATION_S -0.0797680 -0.1415992 -0.0202820    28000 0.0104 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

> H2_MA_MCMC<-MCMCglmm(MA_PERC<-MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION_S, random = ~ RSID3, nitt=30000, burnin = 2000, thin = 1, data = grp_vars)
> summary(H2_MA_MCMC)

```
     Iterations = 2001:30000
     Thinning interval  = 1
     Sample size  = 28000

     DIC: 1531.023
```

G-structure: ~RSID3

```
     post.mean  l-95% CI u-95% CI eff.samp
RSID3 1.023e-06 4.036e-38 2.14e-07    255.5
```

R-structure: ~units

```
     post.mean l-95% CI u-95% CI eff.samp
units    297.2    234.8     359   26602
```

Location effects: MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION_S

```
                post.mean l-95% CI u-95% CI eff.samp    pMCMC
(Intercept)      86.8081  84.2910  89.3597   28000 < 4e-05 ***
POTENCY_S        23.2087  16.9149  29.7115   28318 < 4e-05 ***
RELATIVE_SIZE_S   4.0657  -0.2674   8.3610   27528 0.06793 .
COOPERATION_S    -9.0331 -14.4944  -3.2770   28000 0.00171 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Appendix E

R Syntax and Output for Hypothesis 3 Calculations

Campion Model: Productivity

> H3_model_PPRRF_CAMP<-lmer(PPRRF ~ U1_S + RRF_S + POTENCY_S + WOR
KLOAD_SHARING_S + SOCIAL_SUPPORT_S + COOPERATION_S + (1 | RSID3),
data = grp_vars, REML=FALSE)
> summary(H3_model_PPRRF_CAMP)
Linear mixed model fit by maximum likelihood
t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: PPRRF ~ U1_S + RRF_S + POTENCY_S + WORKLOAD_SHARING
_S + SOCIAL_SUPPORT_S +   COOPERATION_S + (1 | RSID3)
   Data: grp_vars

     AIC   BIC   logLik deviance df.resid
   -85.0   -56.3   51.5 -103.0   170

Scaled residuals:
    Min   1Q Median   3Q   Max
-1.9767 -0.6773 -0.1053 0.5281 3.6277

Random effects:
 Groups   Name   Variance Std.Dev.
 RSID3   (Intercept) 0.00438 0.06618
 Residual   0.02923 0.17097
Number of obs: 179, groups: RSID3, 54

Fixed effects:
                    Estimate Std. Error   df t value Pr(>|t|)
(Intercept)   0.653735   0.016318 40.610000 40.062 < 2e-16 ***
U1_S   0.017818   0.015216 43.650000   1.171   0.2480
RRF_S   -0.009822   0.005488 178.770000 -1.790   0.0752 .
POTENCY_S   0.185913   0.036651 174.400000   5.073 9.97e-07 ***
WORKLOAD_SHARING_S -0.013476   0.028811 174.650000 -0.468   0.6406
SOCIAL_SUPPORT_S   0.076340   0.065055 178.690000   1.173   0.2422
COOPERATION_S   -0.124489   0.059163 178.430000 -2.104   0.0368 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) U1_S   RRF_S POTENC WORKLO SOCIAL
U1_S   -0.218
RRF_S   0.041 0.175
POTENCY_S   0.016 -0.122 -0.107

WORKLOAD_SH 0.021 -0.004 0.263 -0.143
SOCIAL_SUPP 0.015 -0.024 0.134 -0.278 -0.011
COOPERATION -0.018 0.062 -0.109 -0.042 -0.349 -0.784

SVO Model: Productivity
> H3_model_PPRRF_SVO<-lmer(PPRRF ~ U1_S + RRF_S + POTENCY_S + PROP_S
VO.PS + (1 | RSID3), data = grp_vars, REML=FALSE)
> summary(H3_model_PPRRF_SVO)
Linear mixed model fit by maximum likelihood
t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: PPRRF ~ U1_S + RRF_S + POTENCY_S + PROP_SVO.PS + (1 | RSI
D3)
   Data: grp_vars

    AIC   BIC   logLik deviance df.resid
  -83.6   -61.3   48.8   -97.6   172

Scaled residuals:
    Min   1Q Median   3Q   Max
-2.0584 -0.6108 -0.0702 0.5070 3.4273

Random effects:
 Groups   Name   Variance Std.Dev.
 RSID3   (Intercept) 0.004491 0.06701
 Residual   0.030132 0.17358
Number of obs: 179, groups: RSID3, 54

Fixed effects:
            Estimate Std. Error   df t value Pr(>|t|)
(Intercept) 0.736225   0.049529 178.980000 14.865 < 2e-16 ***
U1_S   0.020065   0.015401 45.860000   1.303   0.199
RRF_S   -0.006979   0.005233 177.760000 -1.334   0.184
POTENCY_S   0.149072   0.027971 178.450000   5.330 2.94e-07 ***
PROP_SVO.PS -0.103570   0.058605 171.040000 -1.767   0.079 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) U1_S   RRF_S POTENC
U1_S   -0.124
RRF_S   0.058 0.173
POTENCY_S   0.267 -0.119 0.124
PROP_SVO.PS -0.943 0.054 -0.049 -0.271

Model Comparison: Productivity
> AIC(H3_model_PPRRF_CAMP,H3_model_PPRRF_SVO, k=2)

```
               df       AIC
H3_model_PPRRF_CAMP  9 -84.96401
H3_model_PPRRF_SVO   7 -83.62797
```

Campion Model: Managerial Judgement

> H3_model_MA_CAMP<-lmer(MA_PERC ~ POTENCY_S + WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S + COOPERATION_S + (1 | RSID3), data = grp_vars, REML=FALSE)
> summary(H3_model_MA_CAMP)

```
Linear mixed model fit by maximum likelihood
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: MA_PERC ~ POTENCY_S + WORKLOAD_SHARING_S + SOCIAL
_SUPPORT_S +    COOPERATION_S + (1 | RSID3)
  Data: grp_vars

   AIC     BIC   logLik deviance df.resid
 1537.1  1559.4  -761.5  1523.1     172

Scaled residuals:
   Min    1Q  Median    3Q    Max
-2.5549 -0.6603  0.1031  0.6849  3.0469

Random effects:
 Groups   Name       Variance Std.Dev.
 RSID3    (Intercept)  5.212   2.283
 Residual            285.236  16.889
Number of obs: 179, groups:  RSID3, 54

Fixed effects:
                    Estimate Std. Error      df t value Pr(>|t|)
(Intercept)         86.828     1.305  43.300  66.545  < 2e-16 ***
POTENCY_S           23.317     3.422 178.980   6.815 1.39e-10 ***
WORKLOAD_SHARING_S  -2.591     2.622 179.000  -0.988   0.324
SOCIAL_SUPPORT_S     3.269     6.042 177.870   0.541   0.589
COOPERATION_S       -8.991     5.448 170.330  -1.650   0.101
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) POTENC WORKLO SOCIAL
POTENCY_S   -0.001
WORKLOAD_SH  0.000 -0.138
SOCIAL_SUPP  0.000 -0.285 -0.059
COOPERATION  0.001 -0.036 -0.327 -0.775
```

SVO Model: Managerial Judgement

> H3_model_MA_SVO<-lmer(MA_PERC ~ POTENCY_S + PROP_SVO.PS + (1 | RSI
D3), data = grp_vars, REML = FALSE)
> summary(H3_model_MA_SVO)
Linear mixed model fit by maximum likelihood
t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: MA_PERC ~ POTENCY_S + PROP_SVO.PS + (1 | RSID3)
   Data: grp_vars

    AIC   BIC   logLik deviance df.resid
  1543.1 1559.0 -766.5 1533.1   174

Scaled residuals:
    Min   1Q Median   3Q   Max
 -2.7407 -0.6132 0.1449 0.6502 2.7309

Random effects:
 Groups   Name   Variance Std.Dev.
 RSID3   (Intercept) 2.93   1.712
 Residual   304.12   17.439
Number of obs: 179, groups: RSID3, 54

Fixed effects:
            Estimate Std. Error   df t value Pr(>|t|)
(Intercept) 90.294   4.637 177.070 19.475 < 2e-16 ***
POTENCY_S   17.272   2.600 175.580   6.644 3.69e-10 ***
PROP_SVO.PS -4.383   5.613 178.670 -0.781   0.436

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) POTENC
POTENCY_S   0.249
PROP_SVO.PS -0.958 -0.260
> KRmodcomp(H3_model_MA_SVO,nullmodel_MA)
F-test with Kenward-Roger approximation; computing time: 0.06 sec.
large : MA_PERC ~ POTENCY_S + PROP_SVO.PS + (1 | RSID3)
small : MA_PERC ~ 1 + (1 | RSID3)
        stat   ndf   ddf F.scaling p.value
Ftest 21.554 2.000 174.854   0.99995 4.292e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> KRmodcomp(H3_model_MA_CAMP,nullmodel_MA)
F-test with Kenward-Roger approximation; computing time: 0.06 sec.

large : MA_PERC ~ POTENCY_S + WORKLOAD_SHARING_S + SOCIAL_S
UPPORT_S +
   COOPERATION_S + (1 | RSID3)
small : MA_PERC ~ 1 + (1 | RSID3)
        stat   ndf   ddf F.scaling   p.value
Ftest 13.65   4.00 172.17   0.99985 1.093e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Model Comparison: Managerial Judgement

> AIC(H3_model_MA_SVO, H3_model_MA_CAMP, k=2)
        df   AIC
H3_model_MA_SVO 5 1543.090
H3_model_MA_CAMP 7 1537.087

Secondary Estimations

> H3_model_PPRRF_SVO_MCMC<-MCMCglmm(PPRRF<-PPRRF ~ U1_S + RRF_S
+ POTENCY_S + PROP_SVO.PS, random = ~ RSID3, nitt=30000, burnin = 2000, thin
= 1, data = grp_vars)

> summary(H3_model_PPRRF_SVO_MCMC)

   Iterations = 2001:30000
   Thinning interval = 1
   Sample size = 28000

   DIC: -84.35236

   G-structure: ~RSID3

      post.mean l-95% CI u-95% CI eff.samp
RSID3   0.00138 7.372e-19 0.007286   119.9

   R-structure: ~units

      post.mean l-95% CI u-95% CI eff.samp
units   0.03459   0.0261   0.04305   567.4

   Location effects: PPRRF ~ U1_S + RRF_S + POTENCY_S + PROP_SVO.PS

         post.mean l-95% CI u-95% CI eff.samp pMCMC
(Intercept) 0.739097 0.641049 0.837914   28000 <4e-05 ***
U1_S   0.017356 -0.011099 0.045282   10680 0.2266
RRF_S   -0.006809 -0.017575 0.003409   28000 0.2020

163

POTENCY_S   0.159031 0.100420 0.215899   4539 <4e-05 ***
PROP_SVO.PS -0.106966 -0.228623 0.008752   28000 0.0776 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> H3_model_MA_CAMP_MCMC<-MCMCglmm(MA_PERC<-MA_PERC ~ POTENC
Y_S + WORKLOAD_SHARING_S + SOCIAL_SUPPORT_S + COOPERATION_S, ra
ndom = ~ RSID3, nitt=30000, burnin = 2000, thin = 1, data = grp_vars)
> summary(H3_model_MA_CAMP_MCMC)

 Iterations = 2001:30000
 Thinning interval = 1
 Sample size = 28000

 DIC: 1535.347

 G-structure: ~RSID3

      post.mean l-95% CI u-95% CI eff.samp
RSID3   4.068 3.026e-11   24.21   253.4

 R-structure: ~units

      post.mean l-95% CI u-95% CI eff.samp
units   298.8   237.5   366.7   7753

 Location effects: MA_PERC ~ POTENCY_S + WORKLOAD_SHARING_S +
SOCIAL_SUPPORT_S + COOPERATION_S

           post.mean l-95% CI u-95% CI eff.samp pMCMC
(Intercept)   86.825   84.200   89.413   28000 <4e-05 ***
POTENCY_S   23.297   16.666   30.284   28000 <4e-05 ***
WORKLOAD_SHARING_S   -2.629   -7.843   2.643   21763 0.329
SOCIAL_SUPPORT_S   3.286   -8.919   15.281   25828 0.591
COOPERATION_S   -8.942 -20.248   1.562   25512 0.107

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix F

R Syntax and Output for Hypothesis 4 Calculations

Adding SVO to Best H2 Model: Productivity

> limited_H4_model_PPRRF_SVO<-lmer(PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + PROP_SVO.PS + (1 | RSID3), data = grp_vars)
> summary(limited_H4_model_PPRRF_SVO)
        Linear mixed model fit by REML
        t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
        Formula: PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + PROP_SVO.PS +   (1 | RSID3)
          Data: grp_vars

        REML criterion at convergence: -69.1

        Scaled residuals:
           Min   1Q Median   3Q   Max
        -1.9203 -0.6547 -0.0742 0.5201 3.4988

        Random effects:
         Groups   Name   Variance Std.Dev.
         RSID3   (Intercept) 0.005245 0.07242
         Residual   0.029332 0.17126
        Number of obs: 179, groups: RSID3, 54

        Fixed effects:
                    Estimate Std. Error   df t value Pr(>|t|)
        (Intercept)   0.742773   0.049403 172.970000 15.035 < 2e-16 ***
        U1_S   0.016833   0.015738 43.970000   1.070 0.29065
        RRF_S   -0.009647   0.005295 171.630000 -1.822 0.07022 .
        POTENCY_S   0.210914   0.036238 170.250000   5.820 2.85e-08 ***
        COOPERATION_S -0.083202   0.030313 168.100000 -2.745 0.00671 **
        PROP_SVO.PS   -0.112072   0.058294 164.480000 -1.923 0.05627 .

        Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

        Correlation of Fixed Effects:
               (Intr) U1_S   RRF_S POTENC COOPER
        U1_S   -0.129
        RRF_S   0.045 0.181
        POTENCY_S   0.239 -0.145 -0.025
        COOPERATION -0.054 0.084 0.184 -0.640
        PROP_SVO.PS -0.940 0.058 -0.036 -0.246 0.059

Model Comparison: Productivity

> KRmodcomp(limited_H4_model_PPRRF_SVO, limited_H2_model_PPRRF)

    F-test with Kenward-Roger approximation; computing time: 0.07 sec.
    large : PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + PROP
    _SVO.PS +
      (1 | RSID3)
    small : PPRRF ~ U1_S + RRF_S + POTENCY_S + COOPERATION_S + (1 | R
    SID3)
        stat   ndf   ddf F.scaling p.value
    Ftest 3.6356 1.0000 166.1600   1 0.05828 .

    Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Adding SVO to Best H2 Model: Managerial Judgement

> limited_H4_model_MA_SVO<-lmer(MA_PERC ~ POTENCY_S + RELATIVE_SIZE
_S + COOPERATION_S + PROP_SVO.PS + (1 | RSID3), data = grp_vars)

> summary(limited_H4_model_MA_SVO)

    Linear mixed model fit by REML
    t-tests use Satterthwaite approximations to degrees of freedom ['lmerMod']
    Formula: MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION
    _S + PROP_SVO.PS +   (1 | RSID3)
      Data: grp_vars

    REML criterion at convergence: 1501.6

    Scaled residuals:
      Min   1Q Median   3Q   Max
    -2.4535 -0.5717 0.1129 0.6835 3.0317

    Random effects:
     Groups   Name   Variance Std.Dev.
     RSID3   (Intercept) 11.18   3.344
     Residual   283.31   16.832
    Number of obs: 179, groups: RSID3, 54

    Fixed effects:
            Estimate Std. Error   df t value Pr(>|t|)
    (Intercept)   89.966   4.606 173.440 19.532 < 2e-16 ***
    POTENCY_S   23.829   3.364 173.240 7.084 3.38e-11 ***
    RELATIVE_SIZE_S   3.873   2.236 172.750 1.732 0.08501 .
    COOPERATION_S   -9.290   2.808 173.980 -3.308 0.00114 **
    PROP_SVO.PS   -3.940   5.564 172.210 -0.708 0.47981

    Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

        (Intr) POTENC RELATI COOPER

POTENCY_S   0.235

RELATIVE_SI -0.148 -0.052

COOPERATION -0.052 -0.644 -0.077

PROP_SVO.PS -0.956 -0.246 0.155 0.056

Model Comparison: Managerial Judgement

> KRmodcomp(limited_H4_model_MA_SVO, limited_H2_model_MA)

F-test with Kenward-Roger approximation; computing time: 0.11 sec.

large : MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION_S

+ PROP_SVO.PS +

  (1 | RSID3)

small : MA_PERC ~ POTENCY_S + RELATIVE_SIZE_S + COOPERATION_

S + (1 |

  RSID3)

     stat   ndf   ddf F.scaling p.value

Ftest 0.4914 1.0000 172.2970   1 0.4843

Appendix G

Administrative Reviews

In addition to the University of Louisville Institutional Review Board, the following administrative reviewing agencies have reviewed this work or the research plan as appropriate:

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
SURVEY CONTROL NUMBER: DAPE-ARI-AO-17-19
RCS: MILPC-3
EXPIRES: 04/24/2018

U.S. ARMY HUMAN RESEARCH PROTECTIONS OFFICE
POC: H. Stone
Compliance review certified June 16, 2017

USAREC/3rd Recruiting BDE OPSEC
POC: C. Ebel / G. Smith
Reviewed for public release October 26, 2017

3rd Recruiting BDE PAO
POC: M. Bland
Reviewed for public release October 20, 2017

CURRICULUM VITA

NAME   Gary Daniel Jones

ADDRESS   212 Sage Road
            Louisville, KY 40207

DOB   Pomona, California – October 21, 1977

EDUCATION   B.S., Art, Philosophy & Literature
& TRAINING   United States Military Academy at West Point
                 2000-2004

                 M.S., Organizational Effectiveness, Development & Change
                 University of Delaware
                 2011-2013

AWARDS

PROFESSIONAL SOCIETIES

Member, Academy of Human Resource Development

PUBLICATIONS

Cumberland, D.M., Petrosko, J.M., & Jones, G.D. (accepted). Motivations for pursuing professional certification. *Performance Improvement Quarterly*.

NATIONAL MEETING PRESENTATIONS

Jones, G., Choi, N., & Cumberland, D. M. (2017). The Campion Work Group Effectiveness Model: An Assessment and Revision. Paper presented at the *Academy of Human Resource Development*, San Antonio, TX.

REFEREED JOURNALS

BOOKS AND SYMPOSIA

INVITED PRESENTATIONS